The Alberta Journal of Educational Research

Vol. XIV, No. 1 *March, 1968*

CONTENTS

BOOK REVIEWS

COMMITTEE ON EDUCATIONAL RESEARCH
The University of Alberta

Editorial Comment

This issue of the *Alberta Journal of Educational Research* takes us into the fourteenth year of publication. Readers of long-standing will be aware that the basic concerns of researchers and consumers of research have not changed radically over the years. Writers still concern themselves with the philosophical implications of the choices facing our public schools as technological developments have impact on our society. They are concerned with the validity of our measurement of important attributes and with the soundness of our theoretical assumptions. They report evidence painstakingly gathered to cast light on problems old and new, and they remind us that occasionally we must withdraw from the battle lines to reassess our direction and our methods.

Anderson pleads with educators to begin immediately to study the educational choices made necessary by recent changes in technology and industry. He states that if educators abdicate the responsibility of decision-making they deny themselves the privilege of guiding the choices along philosophically acceptable paths.

In his article, MacKay illustrates the increased efficiency which can result if the scientific approach used in research is applied to the management of that research! Through systematic identification and analysis of the tasks that are critical to the progress of the research, time and money can be invested to greatest advantage.

Sawada concerns himself with the development of a method for assessing "conservation of length" that will result in a relatively pure measure, one virtually independent of verbal abilities. His procedures and reasoning are carefully described so that the reader may judge the applicability of his procedures to other situations. If the investigator considers verbal involvement to be an integral part of a measure of conservation, he will, of course, want a measure that reflects that interrelationship.

Three articles report and analyze data collected in surveys. Rattan and MacArthur present a segment of a longitudinal project that has been reported periodically in the *Alberta Journal of Educational Research.* Their focus is the utility of various culture-oriented and culture-reduced measures of intelligence and achievement in longitudinal prediction for Indian, Eskimo and Metis children. Anant has gathered evidence about the value of two aptitude tests in predicting high school achievement in English and Mathematics. He compares the Differential Aptitude Tests and the Primary Mental Abilities on their predictive validity and efficiency. The relationships between some personality measures and academic achievement are explored in Linton's article. The study he

reports was done with a population of junior high school students from Western Canada.

Greenfield offers a synthesis of leadership research that is not otherwise available. He states the case for a broader conceptualization of leadership which in turn implies a greater concern for theory and for the adequacy of research design. Readers, other than those involved in leadership research, should find that many of the generalizations made by Greenfield also apply to other fields of research.

The statement about preparation of manuscripts, which appears on the inside back cover, indicates a change in editorial policy. The official manual of style has been changed from the *MLA Style Sheet* to the *Publication Manual of the American Psychological Association,* 1967 Revision. The one exception to APA style is that the *Oxford English Dictionary* will continue to be the guide for spelling.

P.A.L.

Alberta Journal of Educational Research
Vol. XIV, No. 1, March, 1968

Two factors which will influence what goes on in schools—technology and the contemporary industrial state—are discussed. The unpleasant consequences of the merging of these influences are depicted and a new remedy suggested.

C. C. ANDERSON
The University of Alberta

Galbraith, Technology, and Education*

There are two increasingly formidable influences on contemporary education. The first is clearly technology;[1] the second, less apparent but probably more profound, is the extent to which the large corporations are controlling the behaviour and beliefs of everyone, especially those of the very young. On this second topic Galbraith[2] has written a rambling, redundant and provocative book which will be considered first.**

Current large corporations such as General Motors, Westinghouse, General Dynamics, Ford, Raytheon, and so forth, have come increasingly to use technology, the ". . . systematic application of scientific or other organized knowledge to practical tasks" (p. 12). This application is possible only if the task, say the construction of an automobile, is broken into subdivisions, each of which is coterminous with some established area of organized knowledge, such as metallurgy, mechanical engineering or inorganic chemistry (p. 12). Some main consequences include the following: the span of time between the initiation and completion of the task increases, specialized experts for each portion of the task must be hired and their work coordinated, the task cannot seriously be changed or the knowledge and equipment related to specialized parts of the task become useless (pp. 13-15), and the necessary commitment of time and capital requires the use of planning and the reciprocal sus-

*The present manuscript was written before the writer became aware of Michael Harrington's "The Social-Industrial Complex", *Harper's*, November, 1967, pp. 55-60, which has the same orientation.

**Page numbers for subsequent references to the Galbraith book will be shown in parentheses following the citation.

pension of the market economy which is dangerously unreliable (pp. 23-26). Only large corporations can carry out such planning (pp. 32-33, p. 74), defined as the ability to organize and employ information, more specifically to foresee and blueprint the actions required between the initiation of a production task and its completion, and to have ready solutions to unscheduled problems which may arise.

This planning must be carried out by a technostructure, a hierarchy of men who bring relevant technical knowledge and experience to group decision-making (p. 59) and who now direct the enterprise in place of the entrepreneur (pp. 89-91). Certain attributes of this technostructural management are worth considering. In the first place, it is ". . . an apparatus for . . . pooling and testing the information provided by numerous individuals to reach decisions that are beyond the knowledge of any one" (p. 77). In this complex and specialized world, no one individual can be relied on to make decisions.

> Forthright and determined administrators frequently react to belief in the superior capacity of individuals for decision by abolishing all committees. They then constitute working parties, task forces, assault teams, or executive groups in order to avoid the one truly disastrous consequence of their action which would be that they should make the decision themselves (p. 65).

The decisions of the technostructure cannot be arbitrarily or capriciously overruled by the intervention of any one individual (pp. 69-70). This is true even of the titular heads of organization, whose main power resides not in their contribution to decision-making but in their choice of the members of groups who make decisions (p. 70).

Secondly, what are the goals that inform the decisions of these men? The first is the preservation of its autonomy by the amassing of capital from earnings (p. 168). Failure to do so would require an appeal to outside sources of capital, an event which would certainly abridge technostructural autonomy, and possibly even necessitate a reckoning with stockholders who are normally about as powerless to influence corporate decisions as the citizens of any community are to change what goes on in schools. Evidence that this autonomy is an important social value can be seen in the restriction of governmental attempts to regulate any industrial enterprise which is inimical to life and limb (pp. 169-170), the imperfect labellings of cigarettes and the failure to restrict the sale of arms being obvious examples. This goal of maintaining a high level of earnings involves the minimizing of risk-taking because prevention of loss is more important than maximum earnings (p. 169), a situation which makes trouble for McClelland's conception[3] now apparently dated, that entrepreneurship involves just such risk-taking.

In order to secure this primary goal of a ". . . secure level of earnings and a maximum rate of growth consistent with the provision of reserves for the requisite investment . . ." (p. 176), the control of the technostruc-

ture must embrace not merely labour and raw materials but also the desires and the aggregate demand of the consumer (p. 200). Consumer demand managed in this way must not be restricted to goods relating to the physical needs of the individual; new and hitherto unknown "psychological needs" must be implanted in the individual by mass advertising. Sales efforts are designed to ". . . shift the locus of decision in the purchase of goods from the consumer where it is beyond control to the firm where it is subject to control" (p. 205). If sales are slipping, corporations can usually find means ". . . to keep exercise of consumer discretion within workable limits" (p. 207).

However, this mass peruasion cannot be expected to deal in subterranean goals which may be at odds with those emphasized by other more obviously attractive, socializing agencies. Accordingly, the direct (professional) and indirect (nonprofessional) advertising efforts of the large corporation are governed by a principle of congruity, which refers to the situation in which the goals of society and the corporation accurately reflect those of the technostructure (p. 161). In Galbraith's words:

> Much of what is believed to be socially important is . . . the adaptation of social attitudes to the goal system of the technostructure. . . . These social goals, though in fact derived from the goals of the technostructure, are believed to have original social purpose (p. 163).

For example, the keeping of aggregate demand reliably high is greatly helped by a large military expenditure by the government, a situation promoting the introduction of advanced technology which would ". . . otherwise be excluded by cost and risk" (p. 229). As Galbraith remarks, "That weaponry in the higher megaton ranges of destructive power has an organic relation to the performance of the economic system leads to unpleasant speculation" (p. 230). However, the extent and effect of this speculation are diminished if the claim is established that the weapons are being manufactured in defense of freedom and not primarily to swell corporation coffers (p. 162). Another important and insidious claim is that the product being manufactured contributes to a high material standard of living (p. 164).

Reviewers of Galbraith's book have largely ignored its significance for professional educators, i.e., the enormous amount of dissonance which this corporate enterprise must create and occasionally attempt to gloss over. Very distressing is the emphasis on consumption of specific and general material products, even though these are more abundant than ever before.

> Morally, we agree that the supply of goods is not a measure of human achievement; in fact, we take for granted that it will be so regarded. . . . Advertising and its related arts thus help develop the kind of man the goals of the industrial system require—one that reliably spends his income and works reliably because he is always in need of more (pp. 209-210).

Another source of dissonance to the wary must come from the deliberate attempt by the technostructure to bamboozle professional experts and the public. For example, Galbraith replaces the Accepted Sequence or flow of instruction from consumer to market to producer by the Revised Sequence, which refers to the situation in which ". . . the producing firm reaches forward to control its markets and on beyond to manage the market behavior and shape the social attitudes of those . . . that it serves" (p. 212). The Accepted Sequence has the big advantage of supposedly ministering to the needs of the individual and partakes of the nature of revealed truth to many. As such it serves as a cloak for the power of the technostructure, even to the extent of sanctioning air and stream pollution, cigarette advertising and industrial squalor generally (p. 217). Imagery has to be continued to maintain the erroneous picture of the corporation's fight for "individual freedom" against the power of government when, in fact, the corporation, often largely and sometimes completely (p. 393) dependent on government contracts, is concerned with advancing its own authority at the expense of the individual.

The second increasing influence on education, technological development, has three connections with Galbraith. Firstly, some of the corporations which are in the defense business have entered the educational market by taking over publishing firms. For example, Raytheon has taken over D. C. Heath and R. C. A. has acquired Random House.[4] The possibility is great, then, that future decisions affecting the curriculum will not be taken by professional educators. Silberman[5] quotes a superintendent as saying that ". . . the center of gravity for educational change is moving from the teacher's college and the superintendent's office to the corporation executive suite." Secondly, the corporation and the school systems generate groups which are characterized by "decision centrality", a situation in which the members of the group search out and transmit factual information to the leader who makes the decision.[6] Because an economical search of this kind must inevitably be carried out by specialists, this situation will eventually resemble Von Bertalanffy's[7] state of "progressive mechanism" in which greater efficiency is attained by having specialized parts the functions of which, however, are not interchangeable. Technological innovations will cause this state, already in evidence with the trend towards specialists and consultants in schools, increasingly to characterize school organization.[8]

What role will the teacher play when, under the euphemistically-labelled computer-assisted (CAI) instruction, he is free from the necessity of supervising rote learning and drill? One of Silberman's answers, that the teacher will do what ". . . only the human being can do, playing the role of catalyst in group discussions . . ."[9] is far from reassuring in that he does not specify what these discussions will be about. Springer[10]

suggests that all personnel under "computerized education" will gain valuable time for the more ". . . creative and substantive aspects of their work", but does not specify exactly what "creative" and "substantive" mean. A third answer begins with an accurate comment by McMurrin[11] to the effect that in schools there is a great neglect of "the inner-life", i.e., ". . . the esthetic pleasures . . . that accrue from sharpening the instruments of sensory perception, or the intrinsic values in the appreciation of poetry and art which are available to those whose education has cultivated their intuitive powers and refined their capacities for sympathy and feeling". Perhaps this instructional (custodial) function, probably the most intractable as far as computerization is concerned, is to be left to the software.

This would be unfortunate for it is here that instruction is at its weakest. Writing from this stance, Friedenberg[12] claims that the school is a means for bringing the individual into line with the rules and behaviour generally regarded as socially desirable by most socializing agencies in any group. Administrators and teachers are specialists in controlling students, and especially dissidents and critics who are not responsive to these rules and behaviour. The general technique for bringing this about is to reinforce the young person's sense of humility and weakness in the face of authority,[13] a task made easier by stressing that accomodative behaviour is a prerequisite for further education and economic advancement.

All sorts of accommodative behaviour are valued by the school; social adroitness,[14] a reliance on external criteria for judging people,[15] a willingness to get along without making trouble for established authority which is, after all, only trying to do its job[16] and a rejection of discrimination which reflects private standards of what is valuable.[17] There is a corresponding distaste for solitary people who pay attention to private feelings and intellectual concerns[18] and who may exhibit, accordingly, unusual and spontaneous responses which are intolerably idiosyncratic.[19] Unfortunately, the study of the fine arts and literature tends to promote just such a subjectivity and individuality. Youngsters who have independent access to these sources of knowledge and experience:

> . . . retain both a standard against which to judge the pattern of values the school conveys and a source of self-esteem beyond its control. But these . . . would also constitute a threat to the mass society if they were allowed to mature. The function of the school, in socializing them, then, is to deprive them of access to that source of self-esteem and to shape their confidence in the standard from which it is derived.[20]

Two remedies are available for this unfortunate state of affairs and for any outbreaks of subjectivity which may or may not be associated with it. The first is to arrange for these subjects to be treated in an authoritative manner by teachers who are quite incapable of having esthetic experience or responding to its communication by others.[21] Special

competence and unique appreciation in this area are to be discredited before the authoritative interpretation of the teacher which ignores alternatives. The second remedy is to rely on the school psychologist to give the same treatment to "the inner-life". This function is to "straighten out" these misguided individuals, a disproportionately larger number of whom come from relatively wealthy homes[22] or are Negroes,[23] by pointing out the present and future advantages of their accommodating themselves to the requirements of the school.[24] Some such custodial function will doubtless be the lot of the future teacher who will be trained as a so-called "psychologist".

Now this sort of schooling is at one with the corporation in discounting the importance of individual decision-making under the guise of protecting his "freedom". Accordingly, the public-relations strategy of both school and corporation must involve the support of high-level "sacred" values which involve no necessary reference to its own business or professional practices and no commitment to any side in current issues. MacIntyre has pointed out, in a different context, that we have specific moral problems like divorce or wars or premarital intercourse and, at the same time:

> . . . a set of exceptionally high-minded, but extremely general moral principles. The difficulty is that we do not know how to connect the principles and the problems. How does one apply such beliefs as that one ought to alleviate suffering or uphold the dignity of the individual to the concrete problems of nuclear weapons or free love? . . . With sufficient ingenuity you can enlist them in support of any and every solution to any and every problem.[25]

This allows a verbal commitment to peace (on earth, good will to men) to sit with a vested interest in the continuance of military spending, in the Cold War and in the beliefs and values consonant with it (pp. 327-333).

If Friedenberg's approach is correct—and certainly teachers' behaviour tends to be of an accommodative sort[26]—a difficult time lies ahead for any group of youngsters who respond to this dissonance by using some of the countermeasures suggested by Galbraith (pp. 345-346), the sceptical scrutiny of official pronouncements[27,28,29,30] and the concentration on nonmaterial social and esthetic values. And yet, if Galbraith is correct, these individuals are the only barrier to a 1984 arranged by corporations and condoned by governments. "The danger to liberty lies in the subordination of belief to the needs of the industrial system. In this the state and the industrial system will be partners" (p. 398). Allowing youngsters to drift into obviously deviant groups which can be pejoratively labelled[31] would be a mark of despair. Conceptual systems theory, as recently advocated by Schroder, Driver, and Streufert[32] may have something constructive to offer.

Schroder, Driver and Streufert[33] distinguish between "content" variables which refer to the dimensions of what a person knows, and "structural" variables, which refer to the number of dimensions along which stimuli are ordered, the number of different "schemata" for organizing dimensions—conjunctive, disjunctive, reciprocal and so forth—and any complexity of that organization. Individuals at a low level of conceptual structure operate in a relatively inflexible way: they use a few compartmentalized dimensions which are coordinated in hierarchical fashion in only one way.[34] Furthermore, as far as conceptual functioning is concerned, stimuli, which are interpreted in terms of one or the other of the dimensions, are distributed dichotomously along the dimension chosen, and new recalcitrant stimuli are either distorted to fit existing dimensions or are gated out.[35] The overt behaviour associated with this simple structure include categorical or black-white judgement,[36] relatively fast decisions in conflict or choice situations because there are few alternatives,[37] and maximal control of behaviour by the stimulous situation. Such an individual is at the mercy of the stimulus, easy game for the advertising agencies.

> The customer . . . must be taught, trained, drilled. By creating is meant the creation of new standards of wants, new appetites, new desires; new dislikes; new standards of taste and propriety; a new environment, physical as well as mental, of work, as well as play.[38]

By contrast, the complex structure, characterized by an increase in the number of dimensions and in the complexity of their coordination, is adaptive in a more flexible way. For example, referring a stimulus to more than one dimension, it exhibits more fine discrimination among stimuli within dimensions, and generates new dimensions and ways in which they can be coordinated so that there is an easy coping with situational change over time or a replacement of a maladaptive interpretation with a potentially more accommodative one.[39]

Schroder, Driver and Streufert point out that the presence of a large number of members with simple structure in a group ". . .implies increasing centralization of power, decreasing interpersonal communication, decreasing flexibility in role assignments . . . ."[40] Stager[41] supported this hypothesis. The speculation here is that there is a reciprocal relationship between simple structure functioning and the centralization of administrative authority, i.e., that the latter may very well be the condition of the former, especially in the case of teachers working under consultants, specialists, and CAI. There is certainly evidence that group structures under stress become hierarchial and members show a regression to simple structure.[42]

This progressive mechanization of individuals and groups may be offset to some extent by education. Schroder, Driver and Streufert[43] contrast a unilateral training environment with an interdependent one.

In the former case, the environment is simplified into ready-made rules which are handed down to the individual and learned by him at the request of the training agent. This has the advantage of placing few demands on the individual, but the disadvantage of discouraging a more self-generated type of information processing. In the case of the second type of training, the agent, by providing information in response to the learner's exploratory questions, encourages the generation of new dimensions and organizing schemata.[44] Presumably teachers and administrators of the latter sort could counteract to some extent the converging pull of technology and corporations in the direction of making us something less than desirably human. MacIntyre has put this general point very well in the context of moral problems:

> Knowing how to approach moral problems is a matter of being the right kind of person. And perhaps the first step towards becoming the right kind of person is to have one's imagination captured by an image of what ought to be and can be.[45]

This is brave talk which, as it stands, must yield to Ausubel's[46] assaults against Bruner's very similar "discovery learning". However, two modifications will place it on a more realistic basis. The first is that its principles must be restricted to material in the fine arts and humanities which lends itself more readily to self-generated rules than that in mathematics and the physical and biological sciences. The second is that this sort of teaching will be restricted to those who can benefit from it, i.e., university-bound students. It is also assumed that the universities will give current teachers in the fine arts and the humanities more than a nodding aquaintance with the social sciences, a point Snow[47] has made in a different context and Holbrook[48] had illustrated, or that a special category of teacher in psychology will be devised whose teaching duties would be concerned with the human situation. A scientific treatment of current behaviour problems—divorce, the effect of drugs, oppression, sexual motivation and so forth—would readily elicit and sustain the interest that high school students ordinarily have in behaviour. For example Platonov[49] has produced a book in which his replies to questions by "young readers" have been developed into a course on behaviour, although the fact that it has a political flavour puts it out of court as far as the present discussion is concerned. At least some experimentation is called for to see what effect such a course would have on students' cognitive structure. The alternatives are despair and hope that a rapprochement between the large corporation and the educational system will be avoided.

REFERENCES

[1] C. E. Silberman, "Technology is Knocking at the Schoolhouse Door", *Fortune*, (Chicago: Times Inc., 1966), pp. 120-125.

[2] J. K. Galbraith, *The New Industrial Estate*, (Boston: Houghton-Mifflin, 1967).

[3] L. Berkowitz, *The Development of Motives and Values in the Child,* (New York: Basic Books, 1964), p. 15.
[4] C. E. Silberman, *op. cit.,* p. 123.
[5] *Ibid.,* p. 120.
[6] H. Hutte, "Decision-Making in a Management Game", *Human Relaitons, 17,* pp. 5-20.
[7] Von Bertalanffy, "General Systems Theory", *Yearbook of the Society for the Advancement of General Systems Theory,* January 14, 1967, *I,* pp. 1-10.
[8] R. F. Campbell, "Tomorrow's Teacher", *Saturday Review,* January 14, 1967, pp. 60-64, p. 73.
[9] C. E. Silberman, *op. cit.,* p. 205.
[10] C. H. Springer, "Technique and Costs: The Systems Approach", *Saturday Review,* January 14, 1967, pp. 56-58.
[11] M. McMurrin, "What Tasks for the School?" *Saturday Review,* January 14, 1967, pp. 40-43.
[12] E. Friedenberg, Coming of Age in America, (New York: Vintage Books, 1967).
[13] *Ibid.,* p. 48.
[14] *Ibid.,* p. 57.
[15] *Ibid.,* p. 65.
[16] *Ibid.,* p. 139.
[17] *Ibid.,* p. 174.
[18] *Ibid.,* p. 66, pp. 211-212.
[19] *Ibid.,* p. 190.
[20] *Ibid.,* p. 182.
[21] *Ibid.,* pp. 184-185.
[22] *Ibid.,* p. 241.
[23] *Ibid.,* p. 197.
[24] *Ibid.,* p. 76.
[25] A. C. MacIntyre, "The Irrelevance of the Church of England", *The Listener,* 1958, *62,* pp. 1054-1055, p. 1058.
[26] C. C. Anderson and S. M. Hunka, "Teacher Evaluation: Some Problems and a Proposal", *Harvard Educational Review,* 1963, *33,* pp. 74-95.
[27] D. Horowitz, *Student,* (New York: Bantam Books, 1962).
[28] H. Draper, *Berkeley: The New Student Revolt,* (New York: Grove Press, 1965).
[29] J. Kozol, *Death at an Early Age,* (Boston: Little Brown, 1967).
[30] C. Murray, "The Bearded, Sockless Radical of Moo V", *The New York Times Magazine,* April, 1967, p. 25, pp. 120-134.
[31] J. Bingham, "The Intelligent Square's Guide to Hippiland", *The New York Times Magazine,* September 24, 1967, p. 25, pp. 67-84.
[32] H. M. Schroder, M. J. Driver and S. Streufert, *Human Information Processing,* (New York: Holt, Rinehart and Winston, 1967).
[33] *Ibid.,* p. 4.
[34] *Ibid.,* p. 15.
[35] *Ibid.,* p. 16.
[36] *Ibid.,* p. 16.
[37] *Ibid.,* p. 16-17.
[38] R. Lewis and R. Stewart, *The Boss: The Life and Times of the British Business Man,* (London: Phoenix House, 1961).
[39] Schroder, *op. cit.,* p. 23.
[40] *Ibid.,* p. 101.
[41] P. Stager, "Conceptual Level as a Composition Variable in Small-Group Decision-Making", *Journal of Personality and Social Psychology,* 1967, *5,* pp. 152, 161.
[42] Schroder, *op. cit.,* p. 101.

[43] *Ibid.*, p. 147.
[44] *Ibid.*, p. 49.
[45] MacIntyre, *op. cit.*, p. 45.
[46] D. P. Ausubel, "Meaningful Reception Learning and the Acquisition of Concepts Analysis of Concept Learning", In H. J. Klausmaier and Chester W. Hains, (Eds), *Analysis of Concept Learning,* (New York: Academic Press, 1966), pp. 157-175.
[47] C. P. Snow, *The Two Cultures: A Second Look,* (New York: Mentor Books, 1963).
[48] D. Holbrook, *The Secret Places,* (London: Methuen, 1964).
[49] K. Platonov, *Psychology as You May Like It,* (Moscow: Progress Publishers, 1965).

Alberta Journal of Educational Research
Vol. XIV, No. 1, March, 1968

The Critical Path Method has been used extensively as an approach to planning and managing large-scale projects. Its application to the planning and implementation of educational research projects seems feasible. A computer program can be used to obtain the various parameters required for implementation of the method. Research project planners and anyone else concerned with sophisticated approaches to planning could make use of this approach.

D. A. MACKAY
The University of Alberta

Application of the Critical Path Method to Educational Research Projects

Introduction

This paper consists of a brief examination of some of the possibilities which sophisticated techniques of project planning may hold for planners of educational research. A very rudimentary example of the application of such a technique to a rather small scale project plan will also be presented. Finally, some attempt will be made to suggest the implication of these techniques for educational research planners.

The Techniques

Of the several newly developed techniques for planning (e.g., Operations Research, etc.), one which has been applied to education, in at least a theoretical way, is PERT (Program Evaluation and Review Technique).[1] In his monograph on the educational applications of PERT, Cook demonstrates, through the use of hypothetical research projects, how that particular technique can be useful in both the planning and the implementation process. Parallel to the development, by the U.S. Navy, of the PERT technique, the Du Pont Company developed a similar technique known as CPM (Critical Path Method).[2] There is

little doubt that the Critical Path Method has become fairly widely diffused in terms of its use in a variety of problem areas. At the same time, one may suggest that some educational researchers will have, at least, heard of the process. However, it appears that application of CPM to educational research project planning is lagging behind the extent of use of PERT. One forms the impression that the U.S. Office of Education has been largely responsible for the diffusion of PERT in that they apparently rely on it as a suitable basis for the development of research projects which they are prepared to support. Whether this lag is more apparent than real is not important for the purposes of this paper. What appears to be useful is a brief exposition of the application of the Critical Path Method, coupled with the observation that PERT is also a powerful tool and that the choice between PERT and CPM would seem to depend on factors not inherent in the techniques themselves.

The Critical Path Method

This Critical Path Method has at least three purposes: (1) to facilitate planning in terms of dividing the project into identifiable "jobs" and scheduling these jobs, (2) to determine which jobs in the project are critical in their effect on the time required for the project, and (3) to determine how resources may be allocated to different jobs in the project in order to meet a target date at minimum cost.

LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4
SURVEY PROJECT	PROBLEM STATEMENT	OBJECTIVES DATA PARADIGMS HYPOTHESES	
	DESIGN	SAMPLE	UNIVERSE DEFINITION SAMPLING PLAN SAMPLE SELECTION
		INSTRUMENTATION	ITEM CONSTRUCTION TRYOUT FORM FINAL FORM
	DATA	COLLECTION	ADMINISTRATIVE PROCEDURES INTERVIEW SELECTION SCHEDULE - ESTABLISHMENT FIELD INTERVIEWS FOLLOW - UP
		ANALYSIS	STATISTICAL TEST INTERPRETATION CODING TABULATION
	DOCUMENTATION	NARRATIVE TABLES CHARTS	

FIGURE 1
TABULAR WORK BREAKDOWN STRUCTURE FOR SURVEY RESEARCH PROJECT

As an aid to logical analysis and planning, CPM can probably best be described by using an example drawn from Cook.[3] Although he was working with PERT, the basic work breakdown and network can, with rather minor modifications, be used to illustrate the Critical Path Method. The example used is really a simulation of a survey research project. The planning process is represented in the analyses depicted in Figures 1-3.

The "work breakdown structure" (Figure 1) has four levels of analysis, increasing in detail from Level 1 to 4. More or fewer levels of detail could have been used depending on how capable the project is of being specified into more and more elementary jobs.

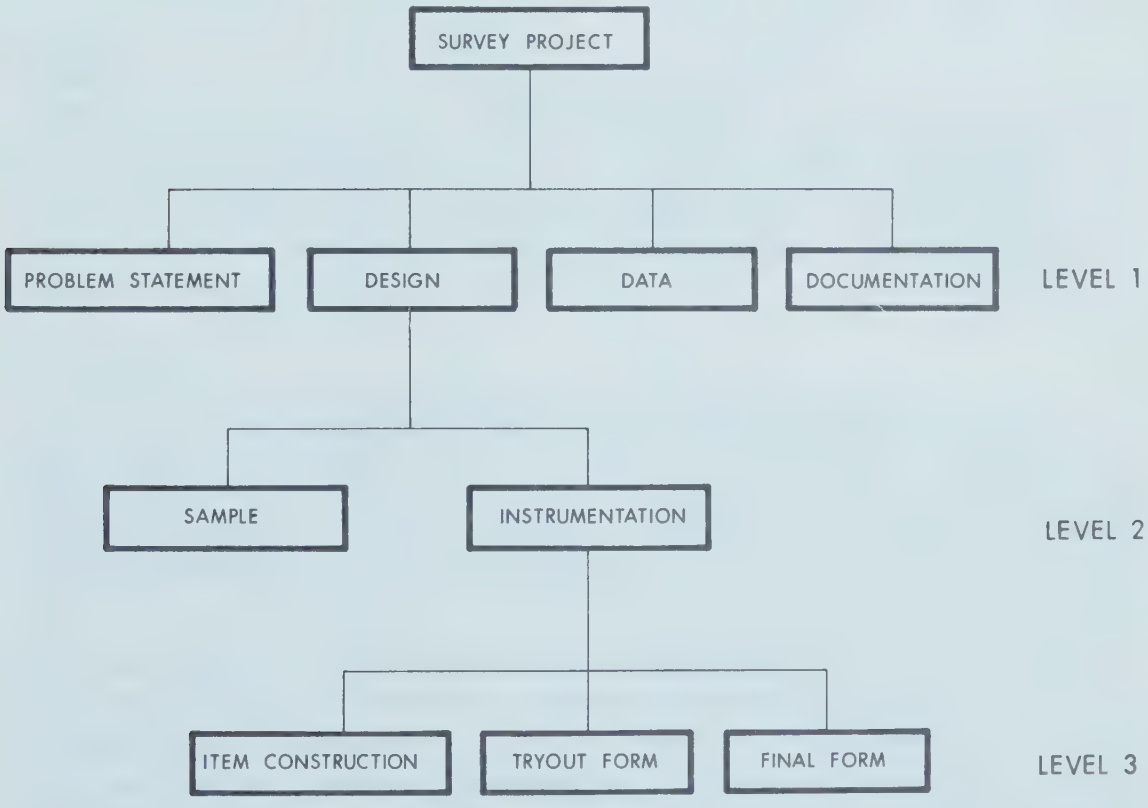
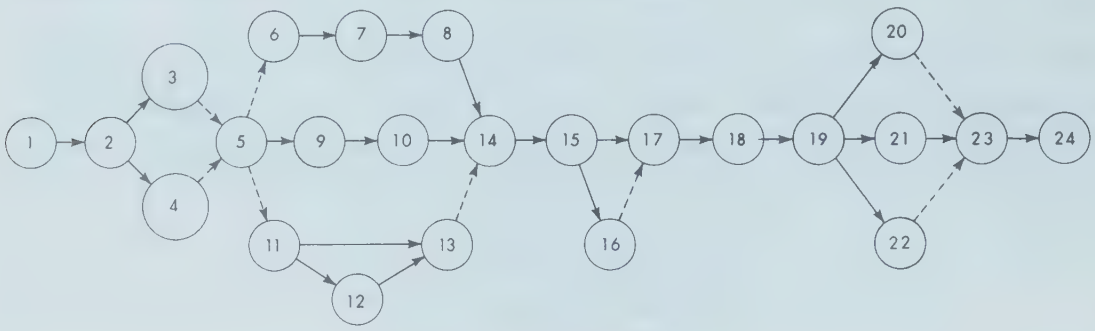


FIGURE 2
PICTORIAL WORK BREAKDOWN STRUCTURE FOR SURVEY RESEARCH PROJECT

The pictorial display in Figure 2 is simply a conventional flow diagram illustrating some of the sequential relations inherent in the project. Finally, in Figure 3 is shown the network for the project. This includes each job in the project and more clearly than either of the other two figures shows the precedence and succession relationships among the jobs. This kind of network is the basic structure to which the critical path algorithm may be applied.

The "critical path" may be defined as the path along and through the summary network which contains only the "critical jobs." These "crit-

ical jobs" are those which directly affect the total project time. If the time taken for these jobs is reduced, then the total time for the project can be reduced. If resources are to be allocated in an attempt to speed up completion of the project, these critical jobs are the ones which merit attention. Attempting to speed up the project by allocating increased resources to all jobs is not necessary.



EVENT IDENTIFICATION

1. PROJECT ORGANIZATION	9. TRYOUT	17. TABULATION
2. OBJECTIVE	10. FINAL FORM	18. STATISTICAL
3. DATA PARADIGM	11. INTERVIEWER SELECTION	19. TESTS SECOND
4. HYPOTHESES	12. ADMINISTRATIVE PROCEDURES	20. INTERPRETATION
5. ITEM CONSTRUCTION	13. SCHEDULES	21. TABLES
6. UNIVERSE DEFINITION	14. FIELD INTERVIEW	22. CHARTS
7. SAMPLING	15. DATA CODING	23. NARRATIVE ONE
8. SAMPLE SELECTION	16. FOLLOW-UP	24. NARRATIVE TWO

FIGURE 3

SUMMARY NETWORK FOR SURVEY RESEARCH PROJECT

As will be seen in a subsequent section of this paper, most of the jobs shown in the example used here were determined to be critical. In more life-like situations where a larger number of jobs is specified, the percentage of critical jobs is, according to one estimate,[4] approximately 10 per cent.

A Computer Program for CPM

The literature on CPM suggests that a good deal of work has gone forward in developing computer programs for various applications of the technique.[5] The program[6] used with the "survey research project" network was one which produces rather limited output, but which is useful as a tool for demonstrating how planning may be facilitated. Before describing the problem which was to be solved, kinds of data required, and some of the output generated, several terms must be defined and the symbols used in this example must be identified.

Definitions and Symbols[7]

Duration (t)—Length of time (in days, weeks, etc.) that job is estimated to take. It is based on knowledge of probable time requirements, learned through one's own or others' experience wiwth the same or similar jobs.

Target (T)—date of completion of the project.

Early Start Time (ES)—For each job, there is an earliest starting time which is the earliest possible time that the job can begin if all its predecessors are also started at their ES.

Latest Start Time (LS)—For each job, there can be defined a late finish (LF) time; this is the latest time that a job can be finished without delaying the project beyond the target date. For each job, therefore, there is a latest start time (LS) which is equal to LF-t.

Early Finish Time (EF)—For each job, the earliest finish time is equal to ES + t.

Free Float Time (FF)—the amount of time a job can be delayed without delaying the early start of any other job (operationally, it is the difference between the EF of the job and the earliest ES of its immediate successors).

Total Float Time (TF)—the difference between a job's early start and its late start. It represents the amount of time a job may be delayed beyond its ES without necessarily delaying the completion of the project. (The critical path, for example, includes those jobs with the smallest TF. All non-critical jobs have greater TF.)

The Problem and Procedures

Problem

The problem which the computer program was capable of solving was essentially as follows: What is the critical path for the survey research network? As sub-problems, the following were raised: What are the early start times, late start times, free float times, and total float times of each job in the network?

Data Required

The input data required were those depicted in the summary network (Figure 3); hence the importance of this network as the basic working data for CPM. Each job was labelled to take account of the sequence of jobs from start to completion of the project. For each job, the number of immediately succeeding jobs and their labels (i.e., the numbers shown in Figure 3) permitted the ordering of the jobs by the computer program in a way which corresponded to the original network. For each job, it was necessary to provide an estimate of the duration. This was done on an arbitrary basis using the "day" as the unit of measure. (In more sophisticated programs, an estimate of the duration can be computed and the probability associated with a weighted estimate can also be obtained.)[8]

It is at this point in the process that human error can probably have its greatest effect. As one of the outcomes of using the technique, however, once the critical path has been determined it is possible to check the time estimates for the critical jobs very carefully rather than attempting to make the time estimate for each and every job in the network perfectly accurate. Then, any changes in time estimates arrived at on the basis of these more careful evaluations can be fed into a new CPM network and a new critical path can be calculated (since the calculations are carried out by a computer program, the iterative nature of this part of the process is not at all a limitation on using the technique).

TABLE I
INPUT DATA

Label	Duration	Name	No. of Following Activities	Labels of Following Activities		
1	1	PROJ. ORGAN	1	2		
2	14	OBJ.	2	3	4	
3	2	DAT. PARAD.	1	5		
4	4	HYP.	1	5		
5	5	ITEM CONST.	3	6	9	11
6	7	UNIV. DEF.	1	7		
7	2	SAMPLING	1	8		
8	2	SAMPLE SEL.	1	14		
9	10	TRYOUT	1	10		
10	4	FINAL FORM	1	14		
11	3	INTERV. SEL.	2	12	13	
12	12	ADM. PROC.	1	13		
13	2	SCHEDULES	1	14		
14	4	FIELD INTERV.	1	15		
15	6	DATA CODING	2	16	17	
16	5	FOLLOW-UP	1	17		
17	5	TAB	1	18		
18	10	STAT. TESTS	1	19		
19	8	TESTS SECOND	3	20	21	22
20	4	INTERP.	1	23		
21	3	TABLES	1	23		
22	2	CHARTS	1	23		
23	2	NARR. ONE	1	24		
24	21	NARR. TWO	0			

A summary of the data input for the problem under discussion here is provided in Table I. For each job, a label (in terms of the sequential numbers in the network), an estimate of duration, the name of the job, the number of following activities, and the label(s) for the "following" activities are input.[9]

Output

In general, the available CPM programs will output a series of "reports" which present the network analysis in terms appropriate to various requirements of the research planner. In the typical programs, these are available as optional output; in the newer, modular programs (such as PMS/360),[10] these various reports are part of the modular structure.

For the survey research project, only one of the "reports" is presented here as Table II. This particular report presents the job in order of early start times. For each job, the input label and duration are shown. The symbols ES, LS, EF, LF, FF, and TF represent the variables defined earlier in this paper. The asterisks beside some of the jobs indicate that these are the "critical jobs" which, therefore, comprise the critical path.

In Figure 4, the network for the survey research project is shown with the partial shading indicating the critical jobs. As was indicated earlier, the proportion of critical jobs is quite high.

TABLE II

JOBS IN ORDER OF INCREASING EARLY START TIMES

LABEL	DURATION		DESCRIPTION	ES	LS	EF	LF	FF	TF
1	1	*	PROJ. ORGAN	0	0	1	1	0	0
2	14	*	OBJ.	1	1	15	15	0	0
3	2		DAT. PARAD.	15	17	17	19	2	2
4	4	*	HYP.	15	15	19	19	0	0
5	5	*	ITEM CONST.	19	19	24	24	0	0
6	7		UNIV. DEF.	24	30	31	37	0	6
9	10		TRYOUT	24	27	34	37	0	3
11	3	*	INTERV. SEL.	24	24	27	27	0	0
12	12	*	ADM. PROC.	27	27	39	39	0	0
7	2		SAMPLING	31	37	33	39	0	6
8	2		SAMPLE SEL.	33	39	35	41	6	6
10	4		FINAL FORM	34	37	38	41	3	3
13	2	*	SCHEDULES	39	39	41	41	0	0
14	4	*	FIELD INTERV.	41	41	45	45	0	0
15	6	*	DATA CODING	45	45	51	51	0	0
16	5	*	FOLLOW-UP	51	51	56	56	0	0
17	5	*	TAB	56	56	61	61	0	0
18	10	*	STAT. TESTS	61	61	71	71	0	0
19	8	*	TESTS SECOND	71	71	79	79	0	0
20	4	*	INTERP.	79	79	83	83	0	0
21	3		TABLES	79	80	82	83	1	1
22	2		CHARTS	79	81	81	83	2	2
23	2	*	NARR. ONE	83	83	85	85	0	0
24	21	*	NARR. TWO	85	85	106	106	0	0

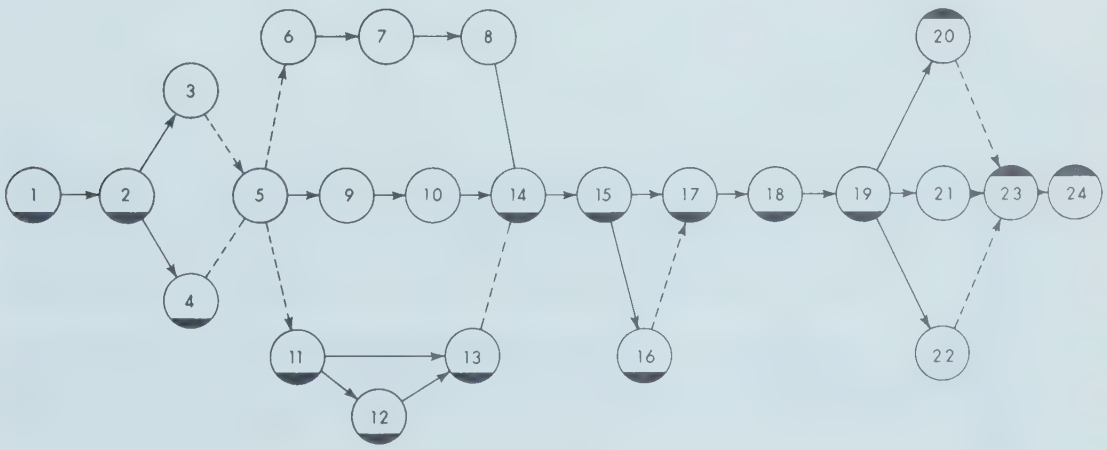


FIGURE 4

THE CRITICAL PATH

Again, it should be emphasized that in a more detailed network the percentage of critical jobs may be as low as 10 per cent.

Summary and Implications

A brief description of the critical path method as an approach to project planning has been provided. A hypothetical example was used to illustrate how a computer program may be used to determine the

critical path and to obtain several useful measures descriptive of a research project.

The implications of CPM (or, for that matter, of PERT) for research project "administrators" would seem to be reasonably clear. First, use of the method requires that systematic planning and a rather detailed job analysis will occur. Second, where time and other resources must be carefully allocated, determining the critical path enables research administrators to make the best use of these resources. Finally, experience in the application of CPM to research project administration is likely to have secondary pay-off. This would be in terms of the skill in using this rather powerful tool in a variety of other applications such as implementation of innovations, training programs for researchers, short and long-term development plans, installation of computer equipment, and so on. Further application of CPM (and its close relatives) to a variety of educational projects of either a research or development nature would seem to merit the interest and activity of researchers.

REFERENCES

[1] Desmond L. Cook, *Program Evaluation and Review Technique Applications in Education,* (Washington, D.C.: U.S. Government Printing Office, 1966).

[2] See, for example, F. K. Levy, G. L. Thompson, and J. D. Wiest, "The ABC's of the Critical Path Method", *Harvard Business Review,* Vol. 41, No. 5, (Sept.-Oct., 1963), pp. 98-108.

[3] Cook, *op. cit.,* pp. 38-41.

[4] Levy *et al., op. cit.,* p. 99.

[5] See, for example, Levy *et al.,* "Mathematical Basis of the Critical Path Method", in J. F. Muth and G. L. Thompson (Eds.), *Industrial Scheduling,* (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1963); I.B.M., Project Management System for IBM System 1360, Technical Paper No. H20-0210-0, (White Plains, N.Y.: IBM, 1966).

[6] J. A. George, *Critical Path Scheduling,* (Edmonton: Department of Computing Science, The University of Alberta, 1966). According to George, this program was adapted from an algorithm developed by C. W. Baker at Stanford University.

[7] These definitions are, in the main, drawn from Levy *et al.;* the symbols are those used in the computer program output.

[8] Cook, *op. cit.,* p. 29.

[9] Further details regarding parameter cards, limits on the number of jobs, etc., are found in George, *op. cit.*

[10] IBM, *op. cit.*

Alberta Journal of Educational Research
Vol. XIV, No. 1, March, 1968

A modified technique for assessing the child's acquisition of length conservation was designed and tested on 62 subjects aged 5-4 to 8-0. The technique attempted to eliminate or minimize the confounding effects of dependence on verbal means of communication, no understanding of instructions, failure to perceive the initial comparison, forgetting, perceptual estimation and guessing.

DAIYO SAWDA and
L. DOYAL NELSON
The University of Alberta

Conservation of Length: Methodological Considerations

Since Piaget's discovery of conservation as a behavioural attribute involved in the child's acquisition of operational thought, many replicative studies have been done confirming the *existence* of the phenomenon of conservation with the result that conservation is generally accepted as a valid attribute in describing the child's cognitive status. These replicative investigations have given way to studies which attempt to relate the phenomenon of conservation to other cognitive variables.[1,2,3,4,5] However, the techniques used in these later investigations to obtain a measure of conservation are essentially the same as those used in the replicative studies. Unfortunately, techniques useful for showing the *existence* of a phenomenon may be neither precise enough nor powerful enough to establish the relationship of the phenomenon to other variables.

Suppose, for example, that we wished to test a simple hypothesis involving the relation of verbal and nonverbal IQ to the acquisition of conservation of length. Suppose further that we classify students as being conservers or nonconservers on the basis of the "usual" procedure typified by the following:

1. Two rods of equal length are placed in front of S as indicated in Figure 1.
   S is asked, "Which stick is longer, or are they both the same length?" Most S will say that they are the "same."

Figure 1.

Rods aligned in "identity position."

2. One of the sticks is transformed (e.g., pushed a few inches to the left or right).
3. S is asked, "Which stick is longer now, or are they still the same length?"
4. Sometimes the question "WHY?" is asked.

In the above procedure there is a heavy reliance on verbal means to communicate to S that he is to respond on the basis of length. In fact, the only way S has of knowing that the attribute under consideration is length is that E used the word "length" or "longer" or some such synonym. Sometimes verbal descriptions, such as "a boy walking along a path" or "two ants going on a walk", are used to communicate to S the criteria he is to use when responding. These verbal descriptions tend to encourage responses based on anthropomorphic criteria which could confound the measure obtained.

If we could couch our description of the criteria of response in a logical rather than a purely verbal form, we could not only give a clearer and more precise definition of the response criteria, but also increase the likelihood that S would respond logically, if, of course, he is capable of doing so. Braine[6,7,8] has reported studies indicating some of the disadvantages of employing a technique which is crucially dependent on verbal means of communicating the response criteria to S. Smedslund[9,10] has cast some doubt on Braine's findings, but regardless of the final outcome of the Braine-Smedslund interchange (see Gruen[11]) it still remains that the measure of conservation derived from the procedure outlined above is so heavily loaded with a built-in verbal factor that its usefulness is severely limited.

Furthermore, such a procedure, dulled as it is by an overlay of verbal dependencies, would probably not possess the necessary sensitivity for detecting the threshold emergence of conservation in young children. In designing training programs for these children it is extremely important that such threshold levels be determined with some degree of precision.

In general, what is needed is a measure of conservation which is maximally free from confounding influences. Such a measure might be called "relatively pure". This paper presents a modified technique for

assessing conservation, and reports, briefly, on an experiment using the technique as a partial evaluation of its effectiveness.

The Problem

In order to design a technique which will give us a relatively pure measure of conservation it would be beneficial to make some basic distinctions concerning the relationship of conservation to concepts in general. At a minimum, we need to distinguish between the concept of a particular attribute (e.g., number, length, time, volume), and the concept of the conservation of the magnitude of that particular attribute. Although the concept of the conservation of the magnitude of an attribute is certainly a part of the concept of the attribute itself, the two are not identical. We are interested only in the conservation aspect of the concept, not in the child's *general* understanding of the attribute.

This basic distinction will be more useful if the following general model for attaining concepts about attributes is adopted. Growth in understanding of a particular attribute can be conceived of as a process of refining a concept already present. With additional refinement, the concept becomes more and more sophisticated. There has to be a starting point, a minimal concept which specifies, rightly or wrongly, when an object possesses the attribute. With respect to the attribute of length, a minimal concept would be that objects take up space, this space being related to the separation of the extremities of the object. More sophisticated concepts would recognize other properties of the expanse between extremities, e.g., that the expanse can change. The recognition of this property would constitute a refinement of the child's concept of length. Other refinements would involve the association of verbal stimuli with the concept. Such words as *big, tall, high, short* and *long* may invoke the concept. The distinction between the words *long* and *high, long* and *tall* and so on would be other refinements. A further refinement would be the recognition of the conditions under which the expanse changes, and perhaps concomitantly, the conditions under which the expanse does not change. The acquisition of these last two refinements is fundamental to conservation since, from an experimental point of view, an operational definition of conservation of length could be in terms of the child's success in specifying which actions (transformations) performed on an object will change the length, and which will leave length invariant. It is the acquisition of the last two refinements that is being tested when testing for conservation of length. Whether or not other refinements have been acquired becomes crucial only if such refinements are necessary for the valid use of the testing technique. The measure of conservation will be confounded with these other refinements if the testing technique is based partially on refinements generally acquired after the refinements being tested for.

The above formulation implies, first, that the testing technique should, as far as possible, be dependent only on refinements which occur prior to the refinements being tested for. The second implication is that in order to test any given refinement, E must communicate clearly to S the basic concept underlying the refinement.

The testing technique described in this paper was designed to fulfill these two conditions. In addition, the design was based on the assumption that children in the process of acquiring conservation of length, being somewhere between Piaget's stages of Pre-Operational and Concrete-Operational Thought, feel most comfortable and operate at their highest levels of competence when dealing with concepts embodied in *physical* objects and in *physical* action.

The Technique

It is readily apparent that the "usual" method of measuring conservation of length does not satisfy the requirements of the implications stated above. First, the method assumes that S has acquired the verbal refinement of ideas specified by the word *length.* However, a given S might associate the word length with the location of the extremities of the objects. He would say that the lengths are the same when objects are aligned in identity position (Fig. 1), but he would judge the length to have changed after the transformation was applied since the location of the extremities did indeed change. Such a S would be classified as a nonconserver. Can we be sure that he is a nonconserver, or might the conservation measure be confounded with the verbal requirements?

Second, E has no way of knowing what criteria S is using to decide whether length has changed. E can ask S why he responded as he did, but such a question assumes again that S has acquired verbal refinements related to his concept of length. The technique reported here attempts to overcome these criticisms as well as some other routine sources of error.

In general, the two criticisms can be overcome by (1) using physical apparatus and physical action to communicate to S the attribute under consideration, and (2) enabling S to give his response in a like manner. The objective with respect to (1) is to present S with a precise definition of the attribute under consideration. When the two rods are aligned in identity position (Fig. 1), the location of the endpoints and the lengths of the rods are the same. Hence the rods themselves do not clearly define the attribute under consideration. In fact, *any* attribute which has the same value for both rods in identity position could be the attribute S chooses to consider.

Since all methods for ascertaining conservation involve the process of comparison (either one object with another, or the same object with itself), and since the process of comparison is essentially the basic process

involved in measurement, the definition of the attribute could be made more precise by making the process of comparison explicit. Since the process of comparison made explicit is simply the process of measurement, the objective can be achieved by *measuring* the magnitude of length, and, as with many other cognitive phenomena, letting our measuring device define that which we are measuring.

Toward this end, one of the rods can be the measuring device, and the other the device to be measured. A crude pair of calipers can be made by attaching physical markers to the ends of the rod and affixing a small handle for ease of operation (Fig. 2). S can then be trained to use the calipers as E wishes him to.

Figure 2.

Two calipers: (left) alone, (right) fitted on an object.

In such a training session, the rods could fit the calipers in only three ways. The rod could be .5 cm. too short for the calipers, just right for the calipers, or .5 cm. too long for the calipers. (Any appropriate expanse could have been chosen.) Thus, the three different fits would be defined to S. Not only would the application of the calipers give a physical definition of length, but the three different fits, differing as they do in only one respect, would also help to communicate to S the attribute under consideration. Therefore, the attribute being manipulated would also be the attribute under consideration. Zimiles[12] has pointed out that many

procedures include so much manipulation of irrelevant variables that S may not be aware of the correct criterion to use in responding. For example, the manipulation of the length of a row of checkers when the attribute under consideration is number may lead S to infer that E really wants him to respond on the basis of length.

The second major objective of the proposed technique is to enable S to use the physical apparatus (1) to facilitate his decision-making about the changes in the magnitude of the attribute, and (2) to communicate his decision to E. The three different fits previously distinguished can be used to accomplish this second objective as well as the first. A three-choice response apparatus could be constructed so that each fit is associated with one of the choices (Fig. 3). The following sequence could then be carried out. A rod and calipers are placed in front of S. S inserts the rod in the calipers (or applies the calipers to the rod) noting the fit. E then removes the calipers and applies a transformation to the rod. S *predicts* how the calipers will fit on the transformed rod by making one of the choices on the response apparatus. Embodied in his choice is his decision about the effect of the transformation on length. Once S understands what is required of him, and this can be accomplished through a training session, there need be no verbal exchange between S and E.

Figure 3.

The response apparatus with rods inserted into calipers by S.

The Experiment

In order to test the feasibility of the proposed technique, and to provide some evaluation of its effectiveness, the technique was tried on young children.

Sample

Sixty-four kindergarten and grade one children were randomly selected from a population of 196 such children attending two schools servicing the Canadian Armed Forces Barracks at Griesbach near Edmonton, Alberta. Most of the fathers were enlisted men of the Corporal or lower rank. Two of the S's had to be excluded from the study: one because he could not meet one of the criteria in the training session and the other because of illness.

Apparatus

There were three major parts to the apparatus: (1) calipers, (2) objects possessing the attribute of length, and (3) a response apparatus.

The calipers were of three magnitudes: 9 cm., 9.5 cm., and 10 cm. The objects were of three kinds—plasticine, wood, and cigarettes—and ranged in length from 1 cm. to 10 cm. The objects could easily be combined to give composite objects possessing "multiple segmented lengths". The response apparatus had 3 doors, each of which could display a candy reward when opened. As suggested earlier, the rods (objects) always fitted the calipers in one of three ways: the rod was .5 cm. too short for the calipers, the rod was just right for the calipers, or the rod was .5 cm. too long for the calipers. The association between each of these fits and the appropriate door on the apparatus was facilitated by placing a "model fit" above the door.

Procedure

Each S was tested individually by the investigator in a session which lasted from 30 to 40 minutes per S. First S was trained and then tested.

The Training Session: As can be gathered from remarks made previously, the crucial part of the technique lies in communicating clearly to S precisely what is required of him and how he is to use the apparatus to make decisions and communicate them to E. Therefore, a detailed account of the training session is given here.

The training session, which lasted from 7 to 12 minutes, had four specific objectives: (1) to train S how to use the calipers, (2) to train S to distinguish among the three kinds of fits, (3) to train S to associate each type of fit with one of the doors on the response apparatus, and (4) to train S to predict the fit of the calipers after E had given the object a transformation.

Objective 1: S was seated in front of the response apparatus (Fig. 3). S was asked if he had ever played with the material on the table and invariably S said no. "Let me show you how." E picked out a 9 cm. caliper and (4, 3, 2)* rods and placed them on the table with one of the rods on end. "We try to push these (pointing to calipers and pushing them slightly toward the rods) onto these sticks." After a slight pause, E said, "But, to make it easier we do this first. We lay the sticks down like this (upright stick is laid horizontally), and put them along here (along the front edge of the box containing the rods) nice and close (gaps are closed between sticks) in a straight line." E brings the sticks back into random formation and says, "Now you try it." Every S could arrange the rods properly. "Now see if you can push these on." S pushes the calipers on and is assisted if necessary. "Those rods fit nicely, don't they? Here, let's try these (9.5 calipers; 5, 3, 2 rods)." Several other examples were given illustrating the three different kinds of fits, thus leading to the next objective.

Objective 2: S was asked if all the sticks fit the same and invariably all S said no. S was encouraged to verbalize the differences in the three fits, but if he did not, he was not urged. If S did volunteer words to describe the fits, then E used S's terminology when referring to the various fits.

Objective 3: After distinguishing the three fits, S's attention was directed toward the three calipers lying on the response apparatus, each of the calipers with an accompanying rod placed parallel to it and about an inch from the calipers. S was directed to place each caliper onto its accompanying rod, and all S agreed that each fit was different. E asked S to tell or show him how they were different, and again E did not insist on verbal replies. S was shown the doors directly below the model fits and how they opened. The one-to-one correspondence between the doors and the fits was established by the designation by E of a fit (either by pointing or by using S's terminology) and the opening of the corresponding door by S.

E then placed calipers and rods in front of S and S applied the calipers. "Which one of these is it like?" If S designated the proper model fit, E asked, "Then which door are you going to open?" If S designated the wrong fit, he was told to look closely and try again. All incorrect designations were the result of hasty and imprecise applications of the calipers, and after trying again these S's realized that they had to apply the calipers carefully if they were to open the "correct" door and find the candy reward. Thus S was aware that close observation of the fit was necessary. S was then asked if he thought there would be candies behind the other two doors, and was invited to try them. There were no candies behind

*Denotes a 4 cm. rod, 3 cm. rod, and a 2 cm. rod.

the wrong doors. S was encouraged to verbalize as to why not, and again E did not insist.

Several other combinations of calipers and rods resulting in different fits were tried, and after S made 4 consecutive correct responses (i.e., found the candy), the final phase of the training session was begun. All except one exceptional S could meet the criterion of 4 consecutive correct responses, and all S took 7 or fewer trials ($p = 4/81$, at a maximum). The exceptional S, who had a very low IQ, was excluded from the study.

Objective 4: At this stage, all S agreed that it was easy to find the candy, as indeed it was. "Well since it's so easy, I'm going to change the game so that it's not quite so easy, and this time you can keep the candy when you find it. I'll show you what we're going to do; watch closely so that you'll be able to find the candy. We're going to do three things. First you put the calipers on the sticks just as before (S does so). Next I'm going to do something to the sticks (E removes the calipers and transforms the sticks). Now *if* you put these (calipers) on again in the way you always do, would they be like this, or this, or this (pointing to model fits)?" Some S pointed at a model fit, but some started to put the calipers on again. For those who pointed at a fit E said, "Where's the candy going to be then? Find it." If S found the candy he was told, "Yes, that's right. That's what it would look like." If S failed to find the candy (i.e., designated the wrong fit) he was told, "No, I guess it wouldn't be that one. It would be like this one or that one." (Parenthetically, it should be noted that up to this point the general objective has been to specify to S that length is the attribute under consideration. It now remains to specify to S that the refinement of conservation is the crucial variable.) For those S who started to put the calipers on the rods, E interrupted and said, "That would make it too easy if I let you do that. I want you to tell me what it would look like without having to put these on. Would it be like this, or this, or this?"

All S were required to do additional examples and all S appeared to grasp what was required of them. In the additional examples E inserted the sentence, "You saw what I did to the sticks. What's it going to look like now, this, or this, or this?" In an attempt to prevent the child from getting the idea that he merely had to choose the original fit again, two of the transformations used changed the length of the rods. The test itself was then begun.

The Test

The 24 test items were given in exactly the same manner as the last examples in the training session. Sixteen of the transformations left length invariant, while eight changed the magnitude of length. The transformations consisted of such actions as (1) rotating the rod(s),

(2) translating the rod(s), (3) adding or subtracting a portion of the rod(s), (4) rolling the plasticine rod, thus elongating it, and (5) cutting a portion of the filter off a cigarette (or adding it on).

Administration of the test: An attempt was made to minimize the confounding influence of several variables:

1. Guessing. In a multiple choice context, guessing plays a prominent role. To counteract its effect, each of the three doors led to a candy reward an equal number of times (8, 8, 8); twenty-four items were used instead of the usual three or four; and S was told that if he was careful and watched what E did to the rods, he could find the candy *every time.*

2. Perceptual estimation. To counteract perceptual estimation the calipers were at no time left in a position parallel to the rods. After removing the calipers, E usually held them partially concealed in his hand, or laid them perpendicular to the rods when possible.

3. Disinterested performance. The use of candy rewards, the novelty of the apparatus and of the technique, along with the active involvement of S throughout the testing session helped to overcome lack of interest.

4. Forgetting. It was crucial that S remember how the rod fit before the transformation was applied. To minimize the role of forgetting, any time that S appeared to forget the initial fit E said, "Did you forget how it fit before? It looked like this one. You saw what I did to the sticks. Which one do you think it's going to be now? Find the candy." Moreover, S was informed that if he ever forgot, he should ask E.

Results

In order to assess the feasibility and effectiveness of the technique it was assumed that if the technique did in fact remove several confounding influences, then the removal of such confounding influences should be reflected in the increased success that younger children would experience with the technique as compared with the success they would have with the usual method. In discussing the age of acquisition of concrete-operational thought processes, Braine[13,14] and Smedslund[15,16] reported their results as threshold ages determined by finding the age interval for which 50% of the subjects in the interval possess the thought process. On this basis they estimated, for Piaget's[17] results, the threshold age for the acquisition of conservation of length to be between 7 and 8 years of age. Murray[18] reported similar threshold ages. Thus, on the basis of the rationale presented here, the effectiveness of the proposed technique can be determined by testing the hypothesis that the technique will result in a significantly lower threshold age for the acquisition of conservation of length.

Before any results concerning the threshold age can be reported, criteria to determine when a given S can be considered a conserver must

be specified. In most tests of conservation, the decision-making process which confronts S can be conceived of as essentially a process of classification. S has to classify the transformations applied to the objects into two mutually exclusive categories: those transformations which leave length invariant and those transformations which change the magnitude of length. In this kind of situation, guessing would give 50-50 results. For the technique used in this study, the involvement of three different fits made possible a finer classification of transformations: those which decreased length, those which left length invariant, and those which increased length. A conserver was one who could classify the transformations correctly in these three categories.

Assuming that all important confounding influences were controlled, and since a three-way response apparatus was used, the probability of a correct response from a nonconserver is no more than one in three. For the purposes of reporting results, the 16 test items which left length invariant were divided into two sets of 8, one involving rotations and the other translations. Each of these sets of 8 was combined with the 8 items which did not leave length invariant. Thus, there were two overlapping subtests of 16 items each. Since half of the items in each subtest changed the magnitude of length, the probability was minimized that any S would maintain the strategy of responding on the basis that the fit after the transformation would be the same as the fit before the transformation.

With $\theta = 1/3$ and $N = 16$, a score of 11 correct responses would be significantly better than chance at the .01 level using the binomial distribution.

Table 1 shows, for each subtest, the number and percentage who reached the criterion in each age interval. Data presented in the table indicate that the threshold age for conservation of length, using the proposed technique, is somewhere between 5-4 and 6-2. These ages are about 2 years lower than those obtained using the usual technique. The hypothesis concerning the effectiveness of the proposed technique is therefore accepted.

Discussion and Summary

A modified technique for assessing the child's acquisition of the conservation of length was presented and evaluated. An attempt was made to control for the following confounding influences: guessing, perceptual estimation, misunderstanding of instructions, misconception of the response criteria, failure to perceive the initial comparison of lengths (perception of the initial "fit"), forgetting the initial comparison, lack of a nonverbal way of expressing decisions, and disinterested performance.*

Perhaps the strongest modification in the proposed technique is the use of physical apparatus to make the technique highly concrete and

*Based on Smedslund's analysis (1963).

more precise, thus overcoming the vagueness and imprecision associated with children's language. However, a price must be paid for introducing such precision, the price being embodied in the logical relations built into the use of the apparatus. The crucial question is whether the dependence on a minimal logical competence is more confounding than the dependence on a minimal verbal competence. Since the experiment resulted in threshold ages about 2 years lower than the usual ages, it is concluded that when the required logic is embodied in concrete operations with physical apparatus, the young child is better able to reveal his threshold acquisition of cognitive processes.

TABLE I

ACQUISITION OF THE CONSERVATION OF LENGTH WITH RESPECT TO AGE

Age Group	No. of Ss	Subtests Serving as Criteria			
		Rotations and Non-Inv		Translations and Non-Inv	
		N	%	N	%
7-2 to 8-0	15	14	93	15	100
6-3 to 7-1	22	19	86	17	77
5-4 to 6-2	25	19	76	16	64

Note: 1. Ss were classified as conservers if the number of correct responses they gave on a particular criterion was significant at the .01 level.
2. "N" designates the number of Ss with conservation.
3. "%" designates the percentage of Ss with conservation.

REFERENCES

[1] A. T. Brace, *The Pre-School Child's Concept of Number,* Unpublished Master's Thesis, The University of Alberta, 1963.

[2] P. C. Dodwell, "Children's Understanding of Number and Related Concepts", *Canadian Journal of Psychology,* 1960, 14, pp. 191-205.

[3] K. E. Feigenbaum, "Task Complexity and IQ as Variables in Piaget's Problem of Conservation", *Child Development,* 1966, 37, pp. 977-983.

[4] I. C. Uzgiris, "'Situational Generality of Conservation", *Child Development,* 1964, 35, pp. 831-841.

[5] E. Halpern, "The Effects of Incompatability Between Perception and Logic in Piaget's Stage of Concrete Operations", *Child Development,* 1965, 36, pp. 491-497.

[6] M. D. S. Braine, "The Ontogeny of Certain Logical Operations: Piaget's Formulation Examined by Nonverbal Methods", *Psychological Monographs,* LXXV, 1959, No. 5.

[7] ..... "Piaget on Reasoning: A Methodological Critique and Alternative Proposals", in *Thought in the Young Child,* William Kessen and Clementina Kuhlman (Eds.), Monograph of the Society for Research in Child Development, 1962, 27, pp. 41-61.

[8] ..... "Development of a Grasp of Transitivity of Length: A Reply to Smedslund", *Child Development,* 1964, 35, pp. 799-810.

[9] J. Smedslund, "Development of Concrete Transitivity of Length in Children", *Child Development,* 1963, 34, pp. 389-405.

[10] ..... "The Development of Transitivity of Length: A Comment on Braine's Reply", *Child Development,* 1965, 36 pp. 577-580.

[11] G. E. Gruen, "Note on Conservation: Methodological and Definitional Considerations", *Child Development,* 1966, 37 pp. 977-983.

[12] H. Zimiles, "A Note on Piaget's Concept of Conservation", *Child Development,* 1963, 34, pp. 691-695.

[13] Braine, *op. cit.,* 1962.

[14] Braine, *op. cit.,* 1964.

[15] Smedslund, *op. cit.,* 1963.

[16] Smedslund, *op. cit.,* 1965.

[17] J. Piaget, *The Child's Conception of Geometry,* (New York: Harper & Row, Publishers, 1960).

[18] F. B. Murray, "Conservation of Illusion-Distorted Lengths and Areas by Primary School Children", *Journal of Educational Psychology,* 1965, 56, pp. 62-66.

Alberta Journal of Educational Research
Vol. XIV, No. 1, March, 1968

For three samples of Metis and Eskimo pupils, several culture-reduced measures of general intellectual ability are shown to have substantial validity in predicting school achievement over three and four year periods. For the Metis sample the predictive validities of two such culture-reduced tests did not differ from those of conventional intelligence tsets.

M. S. RATTAN
Red Deer Junior College

R. S. MacARTHUR
The University of Alberta

Longitudinal Prediction of School Achievement for Metis and Eskimo Pupils

Elley and MacArthur,[1] Olson and MacArthur,[2] and West and MacArthur,[3] have identified several "culture-reduced" tests which seem less bad than certain other widely-used tests in helping to assess the general intellectual potential of Edmonton Whites of lower socioeconomic status, of Edmonton Whites of foreign language background, and of Metis and Indian pupils of the Canadian North. Such tests as the Progressive Matrices, the IPAT test of g, the Safran Culture-Reduced Intelligence Test (SCRIT), and certain subtests of the Lorge-Thorndike Intelligence tests, when compared with conventional intelligence tests, showed less cultural bias for pupils such as these, while at the same time having substantial correlation with more conventional measures of scholastic aptitude and with concurrent school achievement.

Assessment of intellectual potential of pupils is prognostic as well as diagnostic. Given a fair estimate of intellectual potential, we want to predict future academic or general success. The main purpose of the investigation reported here was to study some promising "culture-reduced" tests of general intellectual ability to assess their long-term validity in predicting school achievement for three samples of Indian-Metis and Eskimo pupils. Other problems investigated and reported elsewhere

by Rattan[4] included the stability of the culture-reduced tests and the relationship of both culture-reduced and conventional tests to certain background variables and concurrent school achievement.

Method

Samples

In 1961 a number of ability tests were administered to the 74 Metis pupils who were then in grades one to three at Faust, Alberta. The 1961-65 Metis prediction sample consisted of the 45 of these pupils who were still attending school at Faust and who were tested again in 1965. Mean scores as of 1961 on several variables for the prediction sample of 45 cases were as follows, with corresponding means for the total original sample of 74 cases in brackets: Grade—1.89 (1.82); Age—8.5 (8.6); Socio-economic Index (Blishen Scale)—37.8 (37.5); Language Other Than English Rating—3.3 (3.3); Progressive Matrices I.Q.—102.3 (102.2); California Test of Mental Maturity Total I.Q.—89.9 (88.8); Detroit Beginning I.Q.—89.0 (89.2). Fortunately, on these variables the prediction sample is clearly representative of the original sample.

Some 1965 descriptive data for this Faust Metis 1961-65 prediction sample and for two Eskimo 1962-65 prediction samples from Inuvik and Tuktoyaktuk are reported in Table I.

TABLE I
SOME DESCRIPTIVE DATA FOR THREE PREDICTION SAMPLES

Status in 1965		Faust Metis 1961-65	Eskimo Young 1962-65	Eskimo Middle 1962-65
N		45	62	37
Age	Mean	12-6	11-3	13-11
	Range	9-8 to 15-3	9-3 to 12-3	12-4 to 15-1
Grade	Mean	5.7	3.7	5.8
	Range	3 to 7	1 to 7	3 to 8
SES	Mean	38	38	38
	Range	30 to 57	30 to 57	30 to 57
Sex	M	24	36	18
	F	21	26	19

Procedures

Raw scores on each test or subtest were normalized and converted to T-scores, based on the respective samples taken separately, with means of fifty and standard deviations of ten. Pearson product-moment correlations were then computed between the 1961 or 1962 predictors and the 1965 criterion variables. Differences in correlation with 1965 school achievement for culture-reduced and conventional predictors were tested using the t-test for correlated samples. Multiple correlation coeffici-

ents of two or more predictors with California Total Achievement scores for the Metis sample were also examined, and the relative contributions of the various predictors toward the prediction of school achievement are reported in Rattan.[5]

Results

Some results of the prediction study are reported in Tables II, III and IV. Table II shows predictive validity coefficients between several of the ability tests administerd in 1961 and the Reading, Arithmetic, Language and Total scores on the California Achievement tests administered in 1965 for the Faust Metis prediction sample. The upper half of Table II shows correlations based on scores *normalized* over this group for Coloured Progressive Matrices, Safran Culture-Reduced Intelligence Test, and various subtests of the California Test of Mental Maturity. The lower half of Table II shows correlations based on *intelligence quotients* for Coloured Progressive Matrices, Detroit Beginning First-Grade Intelligence Test, Lorge-Thorndike Intelligence Test, and subtests of the California Test of Mental Maturity.

Table III represents the core of this article. It shows that for the same Metis group, using California Achievement Total in 1965 as the criterion, the so-called culture-reduced measures Coloured Progressive Matrices and SCRIT do not differ significantly from any of the conventional ability

TABLE II

CORRELATION COEFFICIENTS BETWEEN MENTAL ABILITY TESTS ADMINISTERED IN 1961 AND CALIFORNIA ACHIEVEMENT TESTS ADMINISTERED IN 1965

Faust Metis Sample (N=45)

1961 Predictors	1965 Achievement Criteria			
	Calif. Rdg.	Calif. Arith.	Calif. Lang.	Calif. Total
Normalized Scores				
Progressive Matrices	.57	.52	.49	.57
SCRIT	.60	.56	.57	.62
CTMM Spatial	.48	.39	.24	.37
CTMM Non-Language	.59	.54	.37	.50
CTMM Logical	.54	.61	.50	.59
CTMM Numerical	.59	.65	.39	.52
CTMM Verbal	.57	.51	.46	.54
CTMM Language	.70	.74	.56	.68
CTMM Total	.70	.71	.50	.64
Intelligence Quotients				
Prog. Mat. I.Q.	.44	.30	.48	.47
CTMM Non-Lang. I.Q.	.35	.21	.24	.30
Detroit Beg. I.Q.	.37	.37	.43	.42
Lorge-Th. I.Q.	.24	.22	.40	.30
CTMM Lang. I.Q.	.46	.39	.46	.48
CTMM Total I.Q.	.45	.33	.40	.43

TABLE III

SIGNIFICANCE LEVELS OF DIFFERENCES IN CORRELATION COEFFICIENTS WITH 1965 CALIFORNIA TOTAL ACHIEVEMENT SCORES FOR CULTURE-REDUCED AND CONVENTIONAL 1961 PREDICTORS

Faust Metis Sample (N=45)

Normalized Scores	r with Achievement	Prog. Mat. .57	SCRIT .62	CTMM Spat. .37	CTMM N-L .50
CTMM Logical	.59	.....	.....	.....	.....
CTMM Numerical	.52	.....	.....	.....	.....
CTMM Verbal	.54	.....	.....	.....	.....
CTMM Language	.68	.....	.....	**	*
CTMM Total	.64	.....	.....		
Intelligence Quotients		**Prog. Mat. I.Q.**		**CTMM N-L I.Q.**	
	r with Achievement	.47		.30	
Detroit Beg. I.Q.	.42	.....		.....	
Lorge-Th. I.Q.	.30	.....		.....	
CTMM Lang. I.Q.	.48	.....		.....	
CTMM Tot. I.Q.	.43	.....			

..... Not significant at .05 level.
* Significant at .05 level.
** Significant at .01 level.

TABLE IV

CORRELATION COEFFICIENTS BETWEEN MENTAL ABILITY TESTS ADMINISTERED IN 1962 AND VERNON ACHIEVEMENT TESTS ADMINISTERED IN 1965

Eskimo Samples

Sample	1962 Predictors	1965 Achievement Criteria		
		Vernon Eng.	Vernon Arith.	Vernon Vocab.
Esk. Young	Prog. Mat. Coloured	.42	.45	.42
N = 62	Otis Alpha N.V.	.59	.69	.61
Esk. Middle	Prog. Mat. Standard	.62	.69	.49
N = 37	L-Th N.V. Level 3	.71	.71	.66

tests listed down the left in their predictive validity coefficients. From the data of Table II, similar comparisons may be made using California Achievement subtests as criteria.

Table IV reports, for the two Eskimo samples, very useful predictive validity coefficients between culture-reduced ability tests administered in 1962 and Vernon achievement tests administered in 1965. Conventional ability predictors were not available for the Eskimo for comparison purposes.

It should be pointed out that for all three samples no special treatment programs were instituted during the prediction interval, beyond the usual programs of the schools concerned.

Discussion

Most education philosophies require that instruction be adapted to suit the needs and abilities of individual pupils. This presupposes a fair estimate of their intellectual potential. Cronbach[6] suggests that mental ability tests ought to help locate undeveloped potential that novel treatment may bring out. In order that novel treatments may be planned and carried out, there should be reasonable assurance that the tests indicating potential remain fairly stable over periods of time and show substantial long-term predictive validity.

West and MacArthur[7] have provided evidence that, for Metis and Indian pupils at Faust, Alberta and Fort Simpson, N.W.T., such tests as Progressive Matrices, Safran Culture-Reduced Intelligence Test, and Lorge-Thorndike Non-Verbal Intelligence Test Level 3 are significantly less biased against these pupils than are conventional intelligence tests. The current study provides evidence that these three culture-reduced tests also have substantial long-term predictive validity for the Metis and Eskimo samples studied. Further, Coloured Progressive Matrices and SCRIT, for the Metis sample, did not differ significantly from the more conventional California Test of Mental Maturity Language and California Test of Mental Maturity Total in predicting California Achievement Total scores over a four-year period. However, they were inferior to the same conventional tests in the prediction of California Arithmetic scores.

This evidence would seem to suggest that such culture-reduced measures of general intellectual ability should be included in testing programs aimed at helping us understanding and adapt instruction to the abilities and potential of individual native pupils. However conditions are changing so rapidly, and may vary so much from one area to another, that predictive validities may change and vary similarly. Long-term validity studies such as this should also be an integral and continuing part of any testing program.

REFERENCES

[1] W. B. Elley and R. S. MacArthur, "The Standard Progressive Matrices as a Culture-Reduced Measure of General Intellectual Ability", *Alberta Journal of Educational Research,* Vol. VIII, No. 1, 1962, pp. 54-65.

[2] D. R. Olson and R. S. MacArthur, "The Effect of Foreign Language Background on Intelligence Test Performance", *Alberta Journal of Eduactional Research,* Vol. VIII, No. 3, 1962, pp. 157-167.

[3] L. W. West and R. S. MacArthur, "An Evaluation of Selected Intelligence Tests for Two Samples of Metis and Indian Children", *Alberta Journal of Educational Research,* Vol. X, No. 1, 1964, pp. 17-27.

[4] M. S. Rattan, "Predictive Validity and Stability of Measures of Intellectual Potential for Two Samples of Indian-Metis and Eskimo Children", Unpublished Master's Thesis, The University of Alberta (1966).

[5] *Ibid.*

[6] L. J. Cronbach, *Essentials of Psychological Testing,* (New York: Harper, 1961).

[7] *Ibid.*

Alberta Journal of Educational Research
Vol. XIV, No. 1, March, 1968

In this study the author cites experimental evidence to suggest that the use of the Primary Mental Abilities Test in place of the Differential Aptitude Tests can be recommended without loss of predictive validity and with a saving in time for students and counsellors.

SANTOKH S. ANANT
The University of Lethbridge

Relative Effectiveness of Aptitude Tests

Introduction

For many years, schools in Ontario and several other provinces in Canada have been using the Differential Aptitude Tests (DAT) as an aptitude test at the secondary stage. The testing time for all the subtests of DAT is 186 minutes. Even when only four subtests (Verbal Reasoning, Numerical Ability, Abstract Reasoning and Space Relations) are used, the time required for administration of these tests is 115 minutes. On the other hand, testing time for another aptitude test, Thurstone's Primary Mental Abilities, is only 74 minutes, which includes 40 minutes for 'Directions'. In an earlier study by the author, the PMA was found to have a high test-retest reliability over a period of four years ($r_{12} = .78$) and to be a good predictor of school achievement.[1] During 1964-65, the Kitchener Waterloo High School Board asked the author to recommend some aptitude test which could save class-time without any loss in predictive effectiveness. Because of the previous experience of the author with the PMA, a comparative study of PMA and DAT was undertaken. It was hoped that, if PMA could serve as well as the DAT, we would save a great deal of class-time of the students and also make more time available to the guidance counsellors for other duties. The Scholastic Aptitude Tests of Ontario (SATO), administered annually to grade twelve students in the general programme in the Ontario schools, were used as an independent criterion for the study. The SATO is adapted from SCAT, one of the most widely used College Admission tests in the U.S.A.. Because SATO are given toward the end of the school programme and are used

to assess the competence rather than the potential of the student, they can be deemed as achievement tests for the purpose of this study. The grade twelve final percentile marks in English and Mathematics were used as a second criterion, although the subjective nature of these marks was recognized as a limiting factor.

Sample and Methodology

The sample consisted of 373 students in Grade XII of whom 189 were boys and 184 girls. During 1964-65, when the data were collected, they were all enrolled in the general programme in the four collegiates in the Kitchener-Waterloo area. Only those Grade XII students in the general courses who were chosen for the SATO study by the Ontario College of Education were included in the study. These students had already taken SATO and three subtests of DAT in the same year (Grade XII). The only test to be administered was the PMA. The data were analyzed at the Ontario Institute of Education during 1965-66. This report is based on the results of a three-way comparison between PMA, DAT and SATO, and the correlations of these tests with the final marks in English and Mathematics.

TABLE I

COEFFICIENTS OF CORRELATIONS (PEARSON'S r's) BETWEEN VARIOUS SCORES ON PMA, DAT AND SATO

	TOTAL GROUP	BOYS	GIRLS
PMA vs DAT			
V:VR	.41	.36	.45
N:NA	.51	.53	.51
R:AR	.35	.38	.30
S:AR	.38	.36	.19
T:(VR + AR + NA)	.54	.60	.50
T:(VR + AR)	.49	.54	.48
T:(VR + NA)	.55	.55	.57
PMA vs SATO			
V:V	.53	.50	.57
N:Math	.64	.61	.63
T:T	.49	.49	.50
DAT vs SATO			
VR:V	.66	.68	.64
NR:Math	.55	.56	.53
T(VR + NA + AR):T	.63	.66	.59

All r's are significantly different from zero at $p \leq .01$.

PMA SUBTESTS
V = Verbal Meaning
N = Number Facility
R = Reasoning
S = Spatial Relations
T = (Total Score)

DAT SUBTESTS
VR = Verbal Reasoning
NA = Numerical Ability
AR = Abstract Reasoning

SATO SUBTESTS
V = Verbal
Math = Mathematics
T = (Total Score)

Results and Discussion

The correlation (Pearson's Product-Moment r's) between various scores on PMA, DAT and SATO for the total group, and for boys and girls separately, are given in Table I.

Although the correlations between total scores on PMA and DAT are between .50 and .60, the correlations between the subtests in these two tests are rather low. Only verbal and numerical subtests in the two tests seem to have some commonalities. The space subtest of PMA appears to be measuring an ability quite different from the abstract reasoning subtest of DAT. The correlations between PMA subtests and SATO subtests are higher than those between PMA subtests and DAT subtests. It appears that the N of PMA and M of SATO are tapping a common ground. It also appears that DAT and SATO correlate a little better than PMA and SATO on the verbal part while PMA and SATO correlate better than DAT and SATO on the numerical part. The DAT total correlates a little better with SATO total than with the PMA total, though the difference is not statistically significant.

The intercorrelations between subtests of each of the three tests are given in Table II.

TABLE II

COEFFICIENTS OF CORRELATION (PEARSON'S r's) FOR INTERCORRELATION OF SUBTESTS WITHIN EACH OF THREE TESTS—TOTAL GROUP

	V	N	R	S	T	
PMA		.37	.32	.19	.60	V
			.49	.24	.61	N
				.28	.63	R
					.59	S
	VR		NA	AR	T	
DAT			.42	.38	.74	VR
				.39	.67	NA
					.79	AR
	V		M		T	
SATO			.44		.91	V
					.74	M

All r's are significantly different from zero at $p \leq .01$.

While examining the intercorrelations of subtests within various tests it appears that PMA subtests are based on more independent factors than DAT and SATO subtests. All the inter-subtest correlations of PMA with the exception of the one between R and N are lower than the inter-subtest correlations of DAT and SATO. This is an indication of factorially purer subtests as compared with the other two tests. It is also indicative of the superiority of PMA over the other two tests in tapping different abilities of the students. The SATO total seems to be heavily loaded with the verbal part.

A comparison of the intercorrelations of the PMA subtests in this study with corresponding intercorrelations in a previous study by the author are given in Table III.

TABLE III

COEFFICIENTS OF CORRELATION (PEARSON'S r's) FOR INTERCORRELATION OF PMA SUBTESTS IN THE 1961-62 STUDY AND THE PRESENT STUDY

	1961-62 STUDY (139 SUBJECTS)					PRESENT STUDY (373 SUBJECTS)				
	V	N	R	S	T	V	N	R	S	T
V		.29	.58	.35	.64		.37	.32	.19	.60
N			.44	.23	.65			.49	.24	.61
R				.28	.60				.28	.63
S					.45					.59

The correlations of all the subtests with the total are about the same for both studies. Two of the six intercorrelations of the subtests are higher in the 1961-62 study, while three of them are slightly higher in the present study. As PMA's revised version (1962) was used for the present study, the slight differences in the subtest intercorrelations may be caused by the different versions of the test. (The substantial difference [significant at .001 level] in the correlations between R and V might be due to the increased separation between the factors measured by these two subtests.) This provides us with some additional information about the stability of PMA subtests in spite of the recent revision.

The correlations between the Verbal and Numerical subtests of the three tests and the final marks in English and Mathematics are given in Table IV.

TABLE IV

COEFFICIENTS OF CORRELATION (PEARSON'S r's) BETWEEN VERBAL AND NUMERICAL SUBTESTS OF THE THREE TESTS AND FINAL PERCENTILE SCORES IN ENGLISH AND MATHEMATICS

	Tests/School Subjects	English	Math
PMA	Verbal	.33	.17
	Number	.16	.42
DAT	VR	.33	.24
	NA	.24	.39
SATO	V	.42	.27
	M	.25	.50

The correlations of both PMA (V) and DAT (VR) with final scores in English are the same ($r = .33$). The PMA (N) seems to have a little edge over DAT (NA) in its correlation with the final score in Mathematics. SATO (V) and SATO (M) have higer correlations than Verbal and Numerical subtests of PMA and DAT, with final scores in English and Mathematics respectively. This is quite understandable, as SATO

seems to be more of an achievement test than an aptitude test and is expected to correlate higher with final school scores than are the aptitude tests. Correlations in Table IV also support the previous assertion about greater factorial purity of PMA subtests when compared with DAT subtests. Correlations between PMA (V) and the Math. Score and PMA (N) and the English score are lower than the corresponding correlations of DAT and SATO subtests.

Out of the three tests used in the study, PMA is less time-consuming for both students and counsellors. Though the intercorrelations between the total scores on the three tests are not substantially different, the lower inter-subtest correlations of PMA are indicative of its higher factorial purity. While PMA has a certain edge over DAT in its correlation with SATO on the numerical part, DAT is slightly better in the verbal part. There are no significant differences between PMA and DAT subtests in their prediction of final scores in English and Mathematics. Thus, the substitution of DAT by PMA can provide the counsellor with equally valuable information in much less time.

REFERENCES

[1] S. S. Anant and W. A. Ketcham, "Physical Maturation as a Moderator Variable in Predicting High School Achievement from Primary Mental Abilities", *MANAS: A Journal of Scientific Psychology,* 1964, 11, 1, pp. 11-19.

Alberta Journal of Educational Research
Vol. XIV, No. 1, March, 1968

Rokeach's D-scale was administered to 157 Canadian ninth grade students who had been divided into three achievement categories (Failures, Mediocres and Honours) on the basis of a Provincial academic examination. An inverse correlation between achievement level and "dogmatism" was found. The similarities between these results and the results of another similar study using the F-scale on the same general population were discussed.

THOMAS E. LINTON
University of Wisconsin-Milwaukee

Dogmatism, Authoritarianism and Academic Achievement*

Recently this author published a study in which the scores of Canadian junior high school students on the Adorno F-scale were shown to be inversely related to academic achievement.[1] This result was in line with observations of other experimenters.[2, 3, 4]

An instrument similar to the F-scale, the D (Dogmatism) scale, was developed by Rokeach.[5] The D-scale is reported to examine cognitive organization in terms of open vs closed belief systems. A substantial improvement over the F-scale is achieved by Rokeach's test in that it appears not to be affected by the particular political orientation of the testee, as is the case with F. Rokeach documents this claim well. Less well documented, however, is the claim that scores on the D-scale, unlike the F-scale, will not correlate with certain "nuisance" factors such as intelligence or academic ability.[6] The basis for this claim involves the assumption that the items used on the D-scale are not as "obvious" or "direct". Presumably, this factor will not allow the more intelligent individual to "penetrate" the questions by answering in what is thought to be a more socially acceptable manner, as seems to be the case with the F-scale.[7]

*Appreciation is expressed to Mr. P. A. Boudewyns for his assistance in the preparation of this study.

To investigate this factor Ehrlich[8] administered the D-scale to 100 sociology students at the same time that they took a "40 item true-false test of sociology knowledge". He found that dogmatism was inversely related to classroom learning, as evidenced by administering the D-scale and a 40 item quiz at a later point. He also found a —.28 correlation between the D-scale and an aptitude test. Therefore, Ehrlich's study would not seem to uphold Rokeach's assumption noted above.

The present investigation was part of a larger study which sought to examine the personality and social correlates of failure and honours work in a Canadian junior high school. Along with other tests, the D- and the F-scales were administered to a group of 500 high school students. The purpose of this study was to examine Rokeach's claim by investigating the relationship between D-scores and the academic achievement of junior high school students and noting any similarities between these results and the relationships found between the F-scores and achievement, using the same general population.

Method

Three groups of students were administered the D-scale, along with several other tests* and a personal data form. The three groups were selected on the basis of the marks they received on "The Provincial Departmental Examination". This examination is made up by subject matter specialists and is coordinated and administered by the Provincial Department of Education in this Canadian Province.** The examination includes a standardized verbal and quantitative measure of ability and is administered to all ninth grade students in the Province. The results serve as a major determinant of the academic or vocational pattern the student will subsequently follow.

The aggregate mark each student achieves on the examination is placed on a stanine distribution for the total population of students in the Province who take the battery in any one year. Pre-determined cut-off points denote the categories of Failure, Mediocre ("Pass") and Honours. Approximately the top two stanines are designated as Honours, the bottom two as Failures and the middle five as Mediocres. For the present study all of the Failures and Honours and a sample of the fifth stanine (Mediocres) were selected from the total population of public school boys in a large (270,000) western Canadian city.

The procedure consisted of comparing the scores on the D-scale with the results achieved on the Departmental Examinations. Though a four

---

*. . . including McClelland's adaptation of the TAT, Sentence Completion, (Rotter) The California Psychological Inventory, Kluckholn and Strodbeck's "value orientation" measure, and the IPAT Culture Reduced Intelligence Test.

**Alberta.

hour period was permitted for the student testing, most of the Failure group was unable to complete the scale and some of the other instruments that were readily finished by the Mediocre and Honours groups. This absence of closure for the Failures appeared to be due to their lack of interest and motivation in areas related to academic methods, such as filling out a questionnaire and/or taking a test. Therefore, the scores of the Mediocre and Failure groups were collapsed into one category for purposes of statistical analyses. Ultimately only six Failures, 62 Mediocres and 89 Honours were able to complete the scale in the allotted time.

Results

The resultant difference between the D-scale scores for the Failure-Mediocre group and the Honours group was significant ($t\ [155] = 3.19$; $p < .01$). Even though the Failure and Mediocre groups had to be combined, the trend for all three groups was linear and the mean for the Failures was the highest of the three. As the level of achievement rises on the Departmental Examinations, the amount of dogmatism declines, as shown in Fig. 1.

FIG. 1

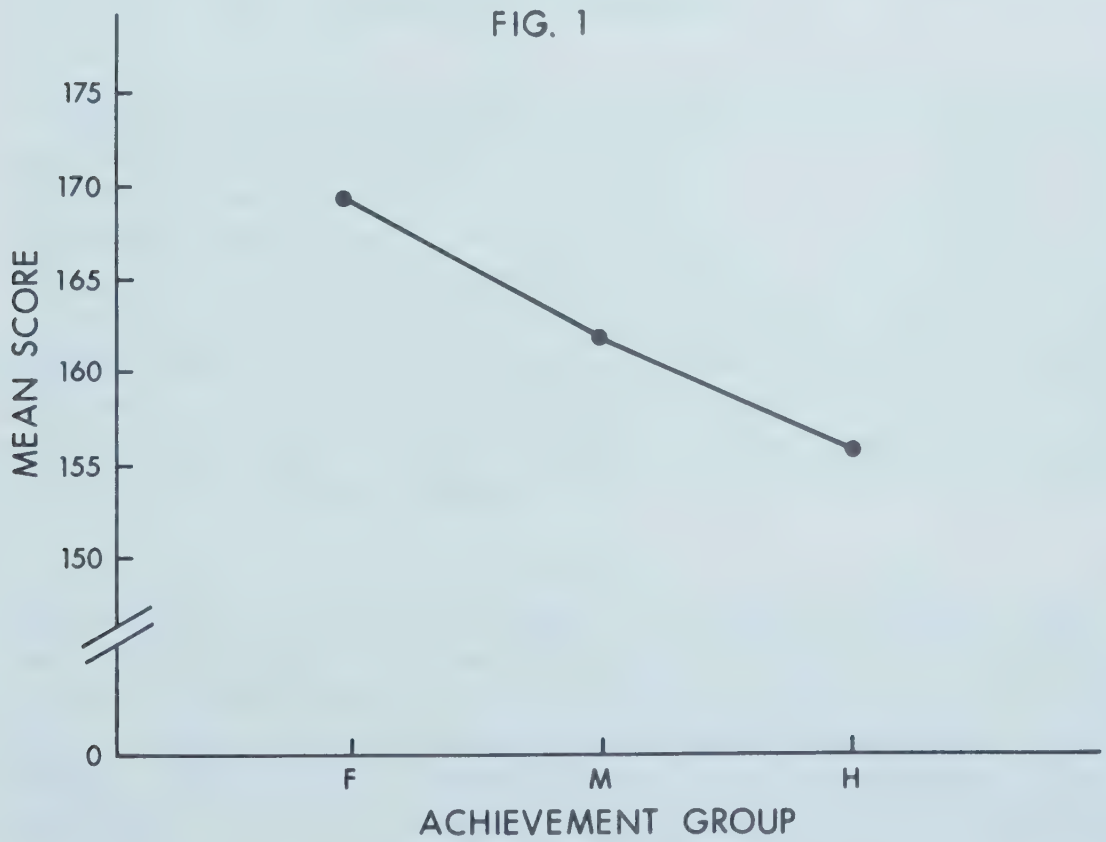
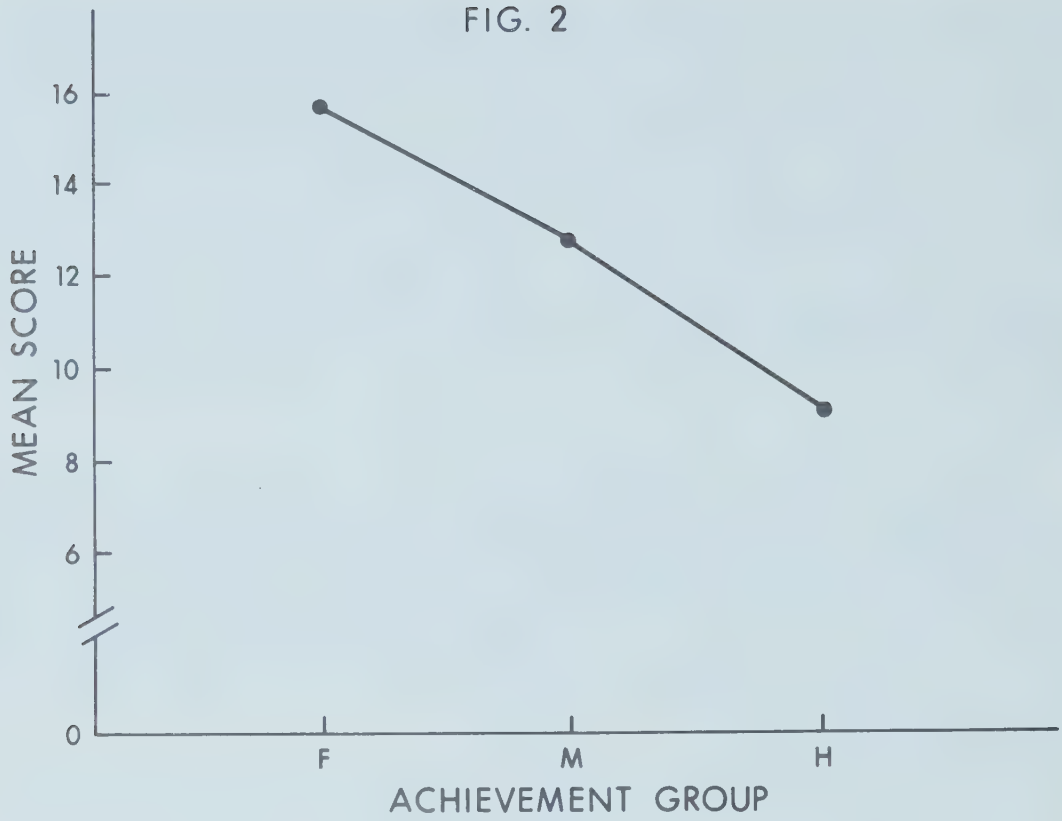


Fig. 2 represents the relationship of the F-scores for the same general population[9] to the relevant achievement groups. Because of the difficulties involved in obtaining closure on these tests from students in

FIG. 2



the lower achievement categories, as mentioned above, no direct correlation was attempted. The observable similarities in the two graphs, however, would seem to demonstrate that academic ability, as defined here, relates to both F and D in a similar manner. The difference between the aggregate F-scores of the combined Failure and Mediocre groups vs the Honours was also significant ($t\ [182] = 5.87$; $p < .001$). It was concluded that the results of this study support Ehrlich's earlier findings that Rokeach's scale is not independent of such factors as academic achievement and intelligence. As with the F-scale, the higher achieving subjects did answer in a more socially desirable manner.

Discussion

One explanation of this result would be that the less well-educated and lower-status students are more dogmatic, and that as education and status rise, the amount of dogmatism (and authoritarianism) tends to decline. This is one of the more usual assumptions about education and social class behaviour. Indeed, in the larger study, noted above, a positive relationship between social class, as measured by income, occupation and education, and the achievement level of the student was found.[10]

It could also be hypothesized that the relationship between achievement and dogmatism was found in this study because the higher achieving student was able to penetrate the instrument and answer in a more

socially acceptable manner. This study was not comprehensive enough to clearly propose either one of these hypotheses.

Further study with the D-scale, using a design similar to the one presented by Cohen (1952), who varied instructional set to measure "penetration" of the F-scale, would be informative.

REFERENCES

[1] T. E. Linton, "Achievement and Authoritarianism as Components of Socio-Cultural Factors in a Canadian Junior High School Setting", *The Alberta Journal of Educational Research,* Vol. XIII, No. 2, June, 1967, pp. 129-134.

[2] Harrison G. Gough, "Studies of Social Intolerance", *Journal of Social Psychology,* 1951, pp. 33, 237-246.

[3] T. S. Cohen, "Is the F-scale Indirect?", *Journal of Abnormal and Social Psychology,* 1952, pp. 47, 732.

[4] A. Davids, "Some Personality and Intellectual Correlates of Intolerance of Ambiguity", *Journal of Abnormal and Social Psychology,* 1955, pp. 51, 415-420.

[5] M. Rokeach, *The Open and Closed Mind,* (New York: Basic Books, Inc., 1960).

[6] *Ibid.,* p. 407.

[7] Cohen, *op. cit.*

[8] H. J. Ehrlich, "Dogmatism and Learning", *Journal of Abnormal and Social Psychology,* 1961, pp. 62, 148-419.

[9] T. E. Linton, "The C.P.I. as a Predictor of Academic Success Among Canadian High School Students", *The Alberta Journal of Educational Research,* Vol. XIII, No. 1, March, 1967, pp. 59-64.

[10] *Ibid.*

Alberta Journal of Educational Research
Vol. XIV, No. 1, March, 1968

The author presents a review and critique of some issues in leadership research. The paper stresses the importance of system-based concepts of organization, examines the Canadian tradition of leader behaviour research as it relates to organizational behaviour, and suggests some theoretical and design problems to be considered in further research.

T. B. GREENFIELD
The Ontario Institute for Studies in Education

Research on the Behaviour of Educational Leaders: Critique of a Tradition[1]

Over the last decade, a number of loosely related studies of the behaviour of educational leaders has been completed in Canadian school systems. These studies are the outgrowth of pioneering work by Hemphill, Stogdill and others,[2,3] now familiarly and widely known as the Ohio Leadership Studies. Even more familiar is the instrument finally evolved as an operational definition of leadership, The Leader Behaviour Description Questionnaire or LBDQ.[4]

Halpin's refinement of the instrument[5] and his research using it to describe the leader behaviour of school superintendents[6] contributed a theoretical framework and empirical tools to educational administrators' longstanding interest in leadership. The fascination of researchers in educational administration with the LBDQ will be commented on later since it is symptomatic of a basic problem apparent in recent research using this instrument. From the Midwest Center at Chicago, LBDQ-based research on the behaviour of educational administrators was taken up by researchers at The University of Alberta and later at other Canadian centres. Much of this Canadian tradition, if we may call ten years of work a tradition, is attributable to the interest and direction of Andrews.[7]

Some Issues in the Study of Leadership

A question frequently asked about leader behaviour is, "How does it relate to individual and group characteristics?" At the present time there are no unequivocal explanations of the relationship between groups and individuals permitting us to claim a satisfactory understanding of leadership phenomena. We do have some agreement about the dimensions of the problem and about the general framework within which the problem may be solved. After a quarter century of research in leadership, it is appropriate to ask where this enquiry has brought us and to refocus upon the underlying issues, namely, the adequacy of theory and research design used in leadership studies.

What is an organization?

The group basis of leadership should cause us to ask what a group is, or, more generally, what an organization is. As educational administrators we are concerned for systems, entities having some formal structure and persisting over time. Theorists like von Bertalanffy and Boulding[8] point out to the concept of system as fundamental to the scientific view of the universe. The importance of this concept is that it provides a framework for handling the problems of growing, living complexities whether these be biological or social. By seeing the system as a process and as a process moreover which acts *as if* it had some goal or purpose, the system concept suggests a mechanism by which groups imprint patterns of behaviour upon the elements comprising them. Systems or, more concretely, role concepts, have been part of the theory of researchers in administrative behaviour for some time, but their implications have, perhaps, not been fully appreciated. Boulding points these out with his typical forcefulness:

> Thus individual electrons come and go in an atom, but once one is captured it must behave in a certain way until it is lost. Similarly atoms come and go in a molecule, but the molecule remains; molecules come and go in a cell, but the cell remains; cells come and go in a body, but the body remains; persons come and go in an organization but the organization remains. What "remains" in the midst of all this flux of components is the "role," the "place," and the relation of roles one to another. A role is a hole, an organization is related and orderly set of holes, and one sometimes catches a fleeting and slightly nightmarish vision of the scientific universe as a set of holes bounded and defined by other holes![9]

Regarding leadership phenomena as explainable in systems and role theory helps to focus attention of researchers upon the context issues in leader behaviour. The Ohio studies and the LBDQ have from the beginning emphasized behaviour and, particularly, patterns of behaviour. Examining the LBDQ as a research instrument, however, one cannot but be struck by the similarity between items on this "test" and the form and content of items from typical interest inventories and personality scales. For example, the following items:

He is friendly and approachable.
He gets along well with the people above him.
He is working hard for a promotion.
He makes his attitudes clear to the group.
He keeps to himself.

The point here is not that these items cannot refer to behaviour *patterns*, but merely that it is easy when dealing with them to think of them as attributes of the indivdual which he brings independently to the group and which determine the leadership functions performed within it. To assume, as suggested by Cattell, that a leader may be defined as "a person who has a demonstrable influence upon group syntality"[10] is surely begging the question when our present knowledge does not enable us to state either the direction or the application of such influence.[11]

This argument suggests that an appropriate framework for the study of leadership must take seriously the organizational context. While making a minimum number of assumptions, adequate theories of leadership should take into account the sources of control in organizations, the dynamics of both internal and external forces which influence the operation of the system. The Ohio studies and the LBDQ itself are built upon descriptions of the kind and frequency of behaviours exhibited by leaders of various groups. The original questionnaire consisted of multiple items defining nine dimensions of leadership as conceived by the original interdisciplinary team of researchers.[12] These items and dimensions were empirically validated by having members of diverse groups describe their leaders in terms of the items. It should be pointed out that these descriptions refer to leader behaviour at a single point in time. Thus while the items may describe behaviour patterns they do so by "photographing" only one point in a process. The items retained in the LBDQ were those differentiating between groups and between dimensions of behaviour. Without questioning the soundness of this technique, it is still necessary to point out that it describes neither changing leadership patterns, the sequence of them, the kind of group in which the patterns were found, nor their relationship to the environment. The position taken in this paper is that questions about dynamics and context are important in the study of leadership. These questions were enunciated some time ago when concerted empirical studies of leadership were begun at Ohio. Since then, however, certain dazzling success with the LBDQ as an instrument to describe leader behaviour has diverted interest from other issues and has allowed assumptions about the individuality of leadership to gain tacit acceptance.

As a reaction to this trend in leadership research, this paper suggests that it is time to re-examine our theories of leadership and to direct attention to some untouched areas of investigation based upon new strategies of enquiry. In particular, it is suggested that research look at

organizational leadership as one of many social phenomena that are shaped by a variety of interpersonal, structural and technical constraints.

Some Dimensions of Organizational Behaviour

From the many issues open to investigation, social psychologists have extracted the individual-group relationship in organizations as the area of prime interest. It is useful then at this point to ask how far small group theorists have taken this investigation and with what results. A recent summary of findings is not encouraging. On the basis of work by Bales, Cartwright, Zander, Homans and others, a recent review article concludes that "there are few occasions when variables at the individual process level are treated as the independent and those at the group level are treated as the dependent."[13] Commenting on this statement and underlining the difficulty of making meaningful statements about group-individual relationships, Yamamoto concludes as follows:

> But, are such effects at all meaningful when the two indices are measuring two facets of the underlying phenomenon at two different levels? It appears rather difficult to establish any simple relationship between the tensile strength of an alloy and the mean value of the molecular weight of the component metals. How then are we to make sense out of the covariation, if any, among the heterogeneous measures of the individual-group-institutional system variables?[14]

This question underlies many group phenomena in education including leadership, authority, and the teaching-learning act itself.

Reference to theorists such as Morris and Seeman,[15] Stogdill,[16] Petrullo and Bass,[17] and Zaleznik and Moment[18] suggests that organizations can best be conceived as multi-faceted systems—systems processing inputs from an environment and yielding outcomes which may modify the organization itself, its components, or the environment in which it exists. Most explicitly and recently Miles[19] has provided a systems model for organizations which relates their functioning not only to leadership, but to such other important group characteristics as morale, communication, change, and adaptation. In building a systems model describing an organization, a useful and perhaps necessary starting point defines four component categories bound together by theoretical connections and to some extent by empirical findings. Without further elaboration on these notions, the categories may be identified as input, social structure, social process and output.

Relationships Among the System Components

Systems are more than a collection of components. The essential properties of a system are the relationships existing between components or between subsystems of components. This condition holds with equal strength for organizations. When we ask, as does much research in leadership, how the characteristics of an individual modify the processes of a

group, we single out for attention only one of the many possible relationships holding among organizational variables. What these relationships are and, more particularly, what causal connections exist within the systems are matters on which research findings are all too scarce.

It would be convenient if some causal chain of relationships held in organizations where social products were a result of inputs from the environment, social process were a result of social structure and where outputs, stemming from process, were fed back again to the environment. In reality organizational relationships are not likely to be either linear or simple; the relationships observed may work in a reverse direction. Some outputs may be determined directly by technology with social process and structure being largely irrelevant to this relationship. Or an observed effect in one subsystem may appear concomitantly with that in another. With a high degree of probability, we may predict conditioning or mediating relationships between the variables. Thus given a particular social structure and process, a technological innovation may produce one effect upon output, though with other social conditions its effects might be different.[20,21]

Research in Leader Behaviour

The Ohio State Leadership Studies, begun in 1945, approached the topic of leadership by "examining and measuring performance or behavior rather than human traits".[22] Because they had little help from previous theory and research, and for other reasons documented elsewhere,[23,24] the Ohio researchers began with a minimum number of assumptions about leadership. Focusing upon how leaders operate, the approach was from the beginning essentially psychometric. Multitudes of items relating to nine supposed dimensions of leader behaviour were contributed by an interdisciplinary team of social scientists. From that point onwards the development of the LBDQ was essentially a process of simplification. The original 1,790 items and nine dimensions were reduced to obtain a comparatively brief, easily administered questionnaire which would provide scales characterizing the behaviour of the leader in a given group while distinguishing it on the same dimension from the behaviour of leaders in other groups. The final coup in this process of reduction was administered by Halpin and Winer[25] who factor analyzed previous forms of the questionnaire to produce the thirty item form composed of the well-known dimensions, Consideration and Initiating Structure.

It is significant that Stogdill, who had been an early worker in the Ohio Studies, later produced an expanded and multidimensional LBDQ-Form XII, claiming, "It has not seemed reasonable to believe that two factors are sufficient to acount for all the observable variance in leader behavior."[26] Although use of the Halpin and Winer LBDQ is frequent in education, it is being replaced by the LBDQ-XII through

Stogdill's own extensive use of it[27] and its increasing applications in education.[28,29,30]

The production of the LBDQ stimulated considerable research on the behaviour of educational leaders with the best known being that of Halpin.[31] Support for these studies came from research and theory of small group behaviour. In this field, structural-functionalist theory was supplanting the situationist approach. The latter approach viewed groups and the leader behaviour required in them as varying in a multitude of situational factors including goals, structure, individual needs and attitudes. The former position held that regardless of the specific group-functions a leader performs, there are basic functions common to all groups and therefore to all leaders. Thus, the LBDQ's purely empirical dimensions of Consideration and Initiating Structure could be interpreted as the leader's efforts to fulfill the group's basic needs for Group Maintenance and Goal Achievement.[32] With this empirical success and theoretical strength behind it, the Canadian tradition of leadership research blossomed by finding various applications of the LBDQ, and later the LBDQ-XII, in school settings.[33]

In reporting this research, the model of an organization as a system will serve to classify the relationships studied. The four-fold model of the organizational system yields ten types of relationships if categories of variables are related two at a time. Thus, a social process variable such as leader behaviour may be related to another social process variable, possibly group characteristics, as measured by Hemphill's Group Dimensions Description Questionnaire.[34] In addition, it could be related to variables in any of the other three categories—inputs, social structure or outputs. The types of relationships possible are indicated in Table I. Multivariate relationships among the variables would greatly increase the number of analyses possible.

TABLE I
TYPES OF RELATIONSHIPS AMONG CATEGORIES OF ORGANIZATIONAL VARIABLES

	Input Variables	Social Structure Variables	Social Process Variables	Output Variables
Input	I/I	I/SS	I/SP	I/O
Social Structure		SS/SS	SS/SP	SS/O
Social Process			SP/SP	SP/O
Output				O/O

The most frequent type of relationship investigated is that between social process and output variables. Specifically, these relationships deal with the consequences of leader behaviour for various measures of output. This emphasis reveals a concern for the outcomes of organizational process that is consistent with common organizational zeal to maximize goal

accomplishment. It is significant, too, that only four of the ten types of relationships have been considered and that multivariate procedures have been used only in a minority of studies. The consequence of this lack of multivariate analysis will be commented upon later. It is sufficient now to say that the emphasis upon leader behaviour-outcome relationship leaves other possible relationships relatively unexplored, particularly the conditioning or mediating kinds of relationships which can be discerned only through multivariate analysis.

Leader Behaviour and Other Social Process Variables

Four studies relate leader behaviour to other social process variables. At the same time that the LBDQ was being developed at Ohio State University, Hemphill[35] developed the Group Dimensions Description Questionnaire (GDDQ) which yields a profile of members' perceptions of group characteristics. The thirteen dimensions included in the questionnaire have been shown by Hemphill to differentiate among teaching groups and a variety of other groups as well. The instrument describes such characteristics as the autonomy, flexibility, hedonic tone, stability and stratification of the groups.

Morris[36] demonstrated that five of the GDDQ dimensions were related to effective principal leader behaviour when effective behaviour is defined as leadership which is jointly and highly considerate and structured. Under these conditions, effective leader behaviour was positively related to participation, polarization and potency in the group; negatively to flexibility and stability. Thus under effective leader behaviour, the teaching groups were found to be characterized by members who gave much time and effort to group activities (participation), who were oriented to goals clear and specific to members (polarization), and for whom the group itself had primary significance (potency). At the same time groups with effective leadership were characterized by members who followed established rather than informal procedures (flexibility), and who felt that the characteristics of the group itself had failed to persist over time (stability). Reverse relationships held among teaching groups where the principals' leadership was described as ineffective. These findings suggest that in activating a group towards clear goals and in building a group that is significant in the eyes of its members, the effective leader lays down firm behaviour expectations and shifts the basic social charteristics of the group.

Thompson[37] investigated the degree to which subordinate supervisory personnel held conflicting expectations for the leader behaviour of superintendents. As with earlier research[38] the subordinates expected more Consideration and Structure than they perceived themselves as receiving from their superintendents. There was also more agreement about the expected leader behaviour than about the perceived, indicating that how-

ever much individuals may agree on the style of leader behaviour they desire to see, they are less agreed about what they *do* see in leaders. Similarly, the subordinates were more agreed about Consideration expected than Structure expected. Thus one of the problems for leaders is meeting the variety of expectations held for their behaviour by subordinates, particularly in the area of the initation of structure.

Two studies of leader behaviour as social process have examined the structure of the research instrument itself.[39,40] Stogdill's claim that two factors, Consideration and Structure, are insufficient to account for all the observable variance in leader behaviour,[41] led him to construct the multi-dimensional LBDQ-XII. The validity of this claim can be tested in part through factor analysis of the instrument. Stogdill earlier observed the high intercorrelations between the sub-scale scores of the LBDQ-XII and performed factor analyses of leader behaviour data obtained in different types of organizations. The results indicated that the LBDQ is indeed reducible to a smaller number of factors, but that the factors emerging depend upon the organization in which the instrument is used.[42] Conflicting evidence is reported on this score by Brown[43] and Punch.[44] This conflict may in part be due to the different approach taken by these investigators. Whereas Stogdill factored the LBDQ-XII data from each organization separately and obtained different factors from each organization (even apparently similar organizations), Brown and Punch factored the data from all schools and then determined how leader behaviour in each school varied in terms of factors identified. Because of the instability of factors based upon small N's, the latter approach seems to be psychometrically more sound.

The factors identified in the work of Brown and Punch are given in Table II. In each of these studies, Factor I composed heavily of Initiating

TABLE II

FACTORS IDENTIFIED IN TWO STUDIES USING THE LBDQ-FORM XII

LBDQ-XII Dimensions	Brown (1967)		Punch (1967)	
	Factor I	Factor II	Factor I	Factor II
Representation	.78[a]		.64	
Demand Reconciliation	.51	.73	.69	.38
Tolerance of Uncertainty		.86		.68
Persuasiveness	.73	.42	.79	
Initiating Structure	.89		.86	
Tolerance of Freedom		.85		.91
Role Assumption	.77	.41	.85	
Consideration		.86	.37	.79
Production Emphasis	.87		.80	—.39
Predictive Accuracy	.62	.63	.81	.34
Integration	.62	.68	.71	.52
Superior Orientation	.57	.50	.55	
Eigenvalue	4.72	4.21	5.21	2.75
% total variation	40%	36%	43%	23%

[a]Loadings smaller than ± .30 have been omitted.

Structure, Production Emphasis and Role Assumption—has been named System Orientation, while Factor II—composed heavily of Consideration, Tolerance of Freedom and Tolerance of Uncertainty—has been named Person Orientation. These two factors bear a significant resemblance to Halpin and Winer's Initiating Structure and Consideration factors in the original LBDQ.

It is not likely that this factoring brings us back to the original dimensions of the LBDQ, since the sub-scales of the LBDQ-XII have instrinsic merit in themselves[45] and help to define the two factors more precisely. At the same time, the methodology of Punch opens the possibility of using determinate factor scores which are psychometrically and practically more valid and useful than the raw sub-scale scores.

Leader Behaviour and Output Variables

The earliest of the studies reviewed here is that by McBeath.[46] This study examined the relationship between the leader behaviour of teachers and their teaching effectiveness as rated by themselves, other teachers in the school, their principals, superintendents and pupils. The study found that rated effectiveness related strongly to teachers' leader behaviour as described by the various judges. The striking exception to this finding occurred when the teachers' descriptions of their own leader behaviour and effectiveness were related. In this case, self-descriptions of leader behaviour and effectiveness had little or no relationship to each other. Feedback to the individual about his own performance was apparently minimal. To the extent that others' ratings of teachers' leader behaviour represent social reality, these findings help to define global ratings of teacher effectiveness, in terms of leader behaviour.

The possibility of halo effects contaminating both leader behaviour and effectiveness ratings is clearly apparent in studies like McBeath's. Halo in the LBDQ was early recognized by its developers,[47] and although McBeath attempted to avoid compounding the effect, it seems likely that both his leadership and effectiveness ratings lie in the same dimension, thus explaining the high relationship between them.

Two studies following McBeath used indices of teacher and school effectiveness which are clearly independent of leader behaviour ratings. Keeler[48] devised an index of school effectiveness based upon pupil achievement and aptitude as measured by a set of external examinations and the Cooperative School and College Ability Tests (SCAT). Using a standardized score representing the difference between actual and predicted pupil achievement, Keeler approximated a measure of pupil growth attributable to the effectiveness of the school in instruction.

With this criterion, Keeler examined the relationships between principal leader behaviour, teacher morale and pupil growth. Both principal leader behaviour and teacher morale were related to pupil growth, though,

contrary to the hypothesis, teacher morale did not act as a mediating variable between principal leadership and pupil growth. Stogdill also failed to establish this hypothesized chain of effect. His contention, based on his own and other research, is "that the executive can more easily influence the cohesiveness and morale than the productivity of the unit that he supervises."[49] Morale and productivity may both be outputs of social process and structure, but any relationship they may have with leader behaviour is likely complex and non-linear.

In Keeler's work, the strength of the relationship between principal leader behaviour and output was considerably weaker than that reported by McBeath between teacher leader behaviour and rated teacher effectiveness. This difference confirms that halo may well have operated in the earlier study and possibly that rated effectiveness might better be regarded as a social process than as an output variable. Another significant finding reported by Keeler indicated a stronger relationship between pupil growth and principal's goal-oriented behaviour than between pupil growth and idiographic behaviour.

Keeler's findings were in part confirmed by Greenfield and Andrews[50] who examined the relationship between teacher leader behaviour and pupil achievement. Instead of the rated effectiveness criterion used by McBeath, they used an index of pupil growth similar to Keeler's. As with McBeath the relationship between teacher leader behaviour and effectiveness was confirmed, though the use of the growth measure reduced the closeness of the relationship considerably. In line with Keeler's findings, too, Greenfield and Andrews reported stronger relationships between growth and structured behaviour than between growth and considerate leader behaviour. They concluded that some twenty per cent of variance in pupil growth was predictable from leader behaviour with most of this (nearly 18%) being due to structure in the teacher's behaviour.

Leader Behaviour and Other Output Variables

Another dimension of output was examined in the work of Miklos.[51,52] Specifically, his study examined the relationship between principal leader behaviour on the one hand and leader ambivalence, group consensus and agreement on the other. Leader ambivalence was defined as the intensity with which the leader holds role expectations for himself; consensus referred to the homogeneity of opinion within the teaching group about the principal's role; agreement referred to the similarity of role expectations between the principal and his staff as a group. Scales defining role expectations were built around Seeman's conceptualization of the leadership role.[53] Contrary to the hypotheses, the leader behaviour dimensions were not consistently related in either linear or curvilinear analysis to the ambivalence, agreement or consensus variables.

Only through multivariate analysis did meaningful relationships emerge between the leader behaviour and the conflict variables defined by Miklos. As might be expected on the basis of system conceptions of organizational behaviour, this finding indicates that many relationships between apparently adjacent and interacting variables are mediated and conditioned by other variables less obviously related to the first pair. For example, Miklos reported that relationships between structure and ambivalence in schools where teachers manifest high consensus were significantly different from those in schools where they manifest low consensus. That is, where teachers were unsure or in disagreement about the principal's role, it was appropriate for the principal to be ambivalent about structure in his leader behaviour. But where consensus existed among teachers about the principal's role his ambivalence about structure was inappropriate. Similarly, agreement between teachers and principals was more likely when principals were described as high in both structure and consideration. Though this finding was not firmly established, it is important even in a tentative form since it questions whether a principal is at the mercy of teachers' expectations or whether "he can in a very real way influence their expectations and bring them close to his self-expectations."[54]

Another finding from the Miklos study bears on this same issue. When tenure of teachers and principals was considered in connection with the structure-consensus relationship, it became clear that this variable mediated the effects of the other two variables. Miklos explained the relationship as follows:

> The longer that principals who are high in structure associate with a particular group of teachers, the more consensus there is within the group of teachers on role expectations for the principal. This does not hold if the principal is low in structure for then the amount of consensus decreases as the length of association increases.[35]

Miklos also reported a similar relationship between tenure, structure and agreement. With structured behaviour on the part of the principal, the agreement between him and teachers increased. Without structure, such agreement apparently decreased over time.

Satisfaction and Confidence as Output Variables

A study by Fast[56] examined the relationship between teacher satisfaction and principal leader behaviour. This study demonstrated that teacher satisfaction was related to the degree of leader behaviour perceived by teachers. Moreover, the greater the discrepancy between leader behaviour expected and perceived, the lower the satisfaction. Expectations of leader behaviour as a single variable had no relationship to teacher satisfaction. It was apparent, too, that consideration perceived was more strongly related to satisfaction than was structure perceived. Furthermore, social and personal relationships among teachers or between

teachers and principals were less important sources of satisfaction than was the professional stimulation which teachers gained from each other and from the principal.

The most recent research relating teacher satisfaction and leadership made use of the LBDQ-XII as an instrument for describing principals' behaviour. Using factor scores derived from the factored LBDQ-XII, Anderson and Brown[57] related principals' leader behaviour to teacher satisfaction, confidence in leadership and ratings of school effectiveness. Most strongly related was leader behaviour and confidence in leadership. While teacher satisfaction showed a similar strong relationship with principal leader behaviour, teachers' ratings of school effectiveness bore no relationship to the principal's leader behaviour. This finding bears out Stogdill's position that group satisfaction and morale as outputs of an organization are not necessarily instruments of productivity. These studies also argue for the desirability of both structure and consideration in leader behaviour, "so long as strength in one form is not cancelled out by a disproportionately poor showing on the other."[58]

The study by Punch[59] investigated the relationship between leader behaviour and social structure in schools. Schools were divided into those where principals' leadership strongly influenced social structure and those where influence was weak. In strong-relationship schools, principals rated their own influence and teacher satisfaction higher than those in weak-relationship schools. Teachers in strong-relationship schools also regarded their principals as more effective than those in weak-relationship schools. School effectiveness ratings bore no relationship to these organizational conditions in schools.

Leader Behaviour and Organizational Climate

A study by Schmidt[60,61] examined the relationship between the LBDQ-XII and Halpin and Croft's Organizational Climate Description Questionnaire (OCDQ).[62] Strikingly, the correlations between these two instruments were rather small. Very few correlations rose above .50 with the highest being .73 between the Production Emphasis scales on the two instruments. In terms of the factors recently identified in the LBDQ-XII,[63,64] the Production Emphasis and Thrust scales of the OCDQ resemble the System Orientation factor while the Esprit dimension resembles the Person Orientation factor. The scales of Consideration, Hindrance, Intimacy and Aloofness on the OCDQ bear virtually no relationship to scales or factors in the LBDQ-XII, suggesting that the factorial structure of the OCDQ is still indeterminate despite contrary claims by its authors.

Leader Behaviour and Input Variables

Comparatively few studies have related input variables; most which have done so have investigated the relationship only as a sub-problem

of the research. Thus, Thompson[65] reported that more highly qualified subordinates expected more structure in their superintendent's behaviour than did less qualified subordinates. More striking findings were brought out by Punch[66] who reported that although schools varied significantly as to their principals' leader behaviour, similar differences were not found between districts. In light of Charters' criticism of the LBDQ,[67] this finding is important in that it indicates the LBDQ-XII is independent of system effects. Punch also reported that school size is unrelated to leader behaviour. Another study of the relationships of a variety of input variables with leadership was made by A. F. Brown.[68] He concluded that none of eight factors such as school bus schedules, density of population and condition of the school building, either together or in combination, significantly affected leadership.

The input variables analyzed in this way included size and type of school, tenure, sex, age and experience of staff and socio-economic status of the school community. These findings give confidence that many of the variables which affect leader behaviour lie within the organization and are not input to the system arbitrarily from the environment. In multivariate analyses some input variables are significant when social structure and leader behaviour are considered jointly. These will be considered below under social structure-process relationships.

Leadership Training as an Input Variable

The only study reported here that considered the effect of leadership training upon leader behaviour is that by Blocksidge.[69] It is also the only study reported which used an experimental design. Twenty-two principals who attended a two-week course in instructional and administrative "leadership" were compared with twenty-two like principals who did not attend the course. Both groups were described as leaders before the course and again after it. Although the course did not give leadership training *per se* and although the leader behaviour descriptions were made in two different academic years after the usual summer break in school terms, significant differences in the behaviour of the two groups were noted. Superintendents and teachers who described the principals' leader behaviour reported an increase in considerate behaviour for the group attending the course. Conversely, the attending principals saw themselves as having less consideration after the course than before it. Thus, the apparent effect of the course was to increase the amount of the principals' considerate behaviour and to give the principals' insights into their own behaviour which caused them to evaluate it less positively. This study suggests other experimental designs for much needed research on changes in leader behaviour over time as other organizational variables are manipulated.

Leader Behaviour and Social Structure Relationships

Some findings about leader behaviour-structure relationships are provided by Thompson[70] and, more significantly, by Punch.[71] Thompson reported that both expectations and perceptions of superintendents' leadership varied by district, though no such variation was found between various groups of subordinates. Thus, different expectations of superintendents' leader behaviour grew up within districts, but not between groups within a given district.

Punch reported findings which make strong inferences about leadership-social structure relationships. Using the Organizational Inventory[72] as a measure of school bureaucratization, Punch found that 57% of variance in school bureaucratization was accounted for by leader behaviour, and offered evidence through multivariate analysis that bureaucratization was dependent upon leader behaviour. In this regard, system-oriented behaviours of principals were positively related ($r = .37$) to bureaucratization in their schools and person-oriented behaviours of these leaders were strongly and negatively related ($r = -.66$) to school bureaucratization.

In schools where principal and teachers had been together for some time, Punch found stronger relationships between leader behaviour and bureaucratization than in schools where togetherness was of shorter duration. Similarly, the relationship between leadership and bureaucratization tended to be stronger when teachers had come to the schools with relatively little experience. Thus, principals seemed able to shape social structure by their leader behaviour when teachers came to a school relatively naive about the social structure and process of schools and stayed in a given school for a long time. It is in this way that schools grow as organizations—as social systems.

Surprisingly, school and system size correlated negatively with bureaucratization. Seventy-five per cent of variance in social structure was accounted for by the most potent variables—system size, school size, togteherness of principal and teachers and, the one most powerful variable, leadership.

A Critique of Theory and Research in Leadership

Examination of leader behaviour studies such as those reviewed above suggest that the time is ripe for researchers to consider a number of critical issues. Greater attention to these issues may mean that research might begin to examine and illuminate some basic leadership problems that have been set aside while other more feasible and pressing problems were being pursued. In conclusion, this paper will consider: (1) the present state of LBDQ-based research, (2) some problems of leadership and group theory, (3) applications of leadership theory, and, (4) research design related to the study of leadership.

The State of Leader Behaviour Research

In reviewing leader behaviour research such as that presented in this paper, the basic similarity of approach in most of them may well strike the reader. While the technical complexity of the analyses and the virtuosity of researchers' statistical skills increase over time, the approach most frequently used relates one variable to another in a rather simplistic fashion. The dependence of these studies upon correlation and factor analytic approaches seldom establishes causal links between variables. By implication, however, the weight of the reported relationships and the repetition of them inevitably suggests determining and mediating effects that are unwarranted by the research design.

These studies have helped to establish a construct validity for the LBDQ in school settings, at least in the two major versions used. The validity was established in terms of a number of other organizational variables, notably output variables of one kind or another. Relationships with other social process variables and with social structure and environmental inputs were also explored. In this regard, however, it is apparent that the emphasis upon leadership-output relationships is perhaps disproportionate. In part this emphasis has served to shift attention from system concepts of organization. Relatively little attention has been focused upon technological influences on organization and upon leader behaviour within it.[73] Similarly, authority relationships between the leader and the environment of his group have received little attention. As Zaleznik and Moment[74] comment, the leader is not "alone" with his group but must interact with higher level leaders whose influence may sometimes be decisive in determining the lower leader's behaviour. We cannot assume that leader behaviour in schools is analogous to leader behaviour in experimental or voluntary small groups or in primary groups.[75]

The research conducted in most of the studies reviewed here may be identified as a variety of "experimental" or "explanatory" survey.[76] This method usually relies upon *ex post facto* manipulation of variables observed at Time 2 to infer causal effects from Time 1. The usefulness of this technique should not be a continuing reason for using it alone when other designs can help to formulate hypotheses that allow the researcher to make "strong inferences" about the direction of observed relationships.[77]

Attention should be drawn, too, to the criterion problem in leadership research. The studies reported here, as already noted, tend to select a rather narrow range of criteria or dependent variables. As Morris and Seeman[78] pointed out some time ago, any organizational variable may become the criterion for the operation of any other. Attention to the problem of how such criteria may best be selected is clearly called

for. The question of new designs for leadership research will be considered in the final section of this paper.

Theory Problems in Group Leadership

The fundamental problem in leadership research may well be the lack of adequate theory or at least the lack of attention to theory. It might be recalled at this point that the Ohio Leadership Studies began virtually without theory, necessarily making a minimum of assumptions about the phenomena studied. Later, theoretical formulations from various sources in organization theory, small group theory and role theory[79,80] were used to provide a framework or model for the phenomena described by the LBDQ. It is well to recall that the developers of the LBDQ saw as one of its prime virtues that it provided quick, easy description of the leader behaviour in groups.[81] The instrument merely describes patterns of behaviours exhibited by leaders at a particular point in time. It is easy now after so many studies based on the LBDQ to confuse such descriptions with theoretical formulations about the nature of leadership. Even the claim that the LBDQ describes patterns of behaviour might be stated more carefully since it is apparent that the descriptions are usually taken at a single point in time, thus providing data that are scarcely adequate to describe complex behaviour processes within an organization. The danger we must now beware is the possibility that researchers may enshrine a single instrument as an adequate model of group process.

Clearly, more dynamic statements about leadership processes are needed than are implied in the LBDQ alone. As a beginning Homans' conceptions of leadership[82] may serve an important purpose. Unlike LBDQ-bound statements about leadership, Homans' concepts emphasize the process and the system; he characterizes the leader as the person in the group moving it from one steady state to another. His "rules" about leadership deal with active leading, with role maintenance, with authority, obedience and channels of communication. Thus, Homans' conceptions of the leadership function contrast rather sharply with the psychologically and psychometrically oriented items of the LBDQ. As Gouldner suggests, evidence of the traits approach to leadership in LBDQ-based research may be inevitable, "peering through the cracks" of an apparently system oriented approach.

Other useful approaches are suggested in the work of organizational analysts such as Likert,[83] Kahn and Katz,[84] and Zaleznik and Moment,[85] where leadership is examined in terms of the general supervisory relationship and other variables often not considered in typical LBDQ-based research. Recent work by McNamara and Enns[86] moves in this direction by examining school leaders in terms of theory and variables developed by Fiedler.

Leadership is clearly a phenomenon of social groups of many kinds. Studies of leadership in education-oriented groups cannot usefully progress in isolation from studies in other kinds of groups. Stogdill's contention that leader behaviour factors are a function of the particular group clearly conflicts with factor analysis of leader behaviour in school groups. This issue, already identified in this paper, offers an opportunity to examine the source of conflict and to develop research whose findings could have important implications for building leadership theory.

More theoretical problems lie in the basic concepts of group that have been used in leadership research. The relationship between individual properties and group properties is still unclear. Underlying this problem is the question of what variables adequately express the syntality of the group rather than merely the average value of heterogeneous individual characteristics.[87] It is apparent that the LBDQ sums and averages individual perceptions and judgements, thus raising the problem of halo and the difficulty of relating phenomena at one level to those at another. Yamamoto's analysis of this problem has already been referred to.[88] Since this is the same problem now receiving attention in research on teacher-pupil interaction,[89] work in this field might profitably be related to studies in leadership.

The question of what a "group" is in a large school offering a highly diversified curriculum to hundreds or thousands of pupils is a question leader behaviour researchers in education have largely avoided. One means of avoiding it has been to study only relatively small schools, usually at the elementary grade level. It is now time that this question was attacked more directly in view of its theoretical and practical importance. Finally, in the area of theoretical problems, one might mention the question of how leaders exert influence in schools and what mechanisms and resources are involved in this process.[90] Zaleznik and Moment[91] mention the process of group reality testing as an important aspect of leadership processes. Incorporation of these and other novel variables in research on educational leaders is surely now desirable and possible.

Applications of Leadership Research

Applications of the findings of research in leadership to educational practice have been relatively few in number. Unlike the practice in industry,[92] education has not used the LBDQ as an instrument for controlled attempts to train leader behaviours. This area of research clearly warrants development as this line of investigation has yielded important findings from its use in business settings.[93] Indeed, the LBDQ has been used in leadership clinics for school administrators[94] and has provided important data for research; these clinics have also given practitioners insights into leadership theory and feedback about their own leader

behaviour that would otherwise be inaccessible to them. Other more extensive clinical applications and research are needed. Explaining leadership theory to a principal and describing his behaviour by way of the LBDQ is roughly equivalent to explaining theories of intelligence or personality to a client before having him complete the Stanford-Binet or the MMPI. Without help from his therapist, what can the client do with his scores from such tests? And what can a leader do about his "scores" on the LBDQ or other similar "tests"? Clearly, giving the test is only the beginning of clinical procedures. Theories of leadership are not theories of leading any more than theories of learning are theories of teaching. In studies of leadership we are only beginning to understand the nature of the phenomenon. How this knowledge can be used to improve leader behaviour and organizational performance has largely still to be discovered.

Research Design for Leadership Studies

The final issue to be raised here concerns research designs appropriate for the study of leader behaviour in organizations. The primary need in leadership research is to create designs which will permit the study of process. The LBDQ is designed to slice through an organization and describe conditions at a fixed point in time. This procedure does not give much information about process. At least, repeated observations using the LBDQ in the same organization are required. Laboratory groups and participant observer techniques could also be used to advantage in connection with leadership research. To date, very few studies in educational leadership have used such procedures.

Initiated structure and manifest consideration have been the data collected for most leader behaviour studies. Questions about the establishment of structure or the development of consideration have not been and cannot be answered in current research designs. What is needed is an understanding of the dynamics of group process. We need to know, for example, how changes in interaction, amount of structure or goal definition affect the attitudes, cohesion and productivity of teaching groups. Evidence about such matters has been obtained from experimentally contrived small groups. However, such studies have not usually attempted to simulate teaching groups, and virtually no studies have examined this kind of problem in teaching groups actually functioning in schools.

Stogdill describes group processes and leadership as observed in experimentally contrived groups:

> . . . tension and conflict are involved in the development of structure in experimentally created groups given the task of discussing a social problem and preparing a written solution. Initially, the members exhibit a high degree of task-oriented performance, but this preoccupation with the task decreases in later sessions. Comparatively small amounts of disagreement

> and tension are exhibited in the first session, but these show a sharp increase in the second session and thereafter decline. Agreement declines sharply from an initial high level, but this is compensated for by an equally sharp rise in solidarity and tension release.[95]

Statements of this kind cannot be made on the basis of current research on the behaviour of school leaders. The reason for this is largely due to the one-shot designs used in leader behaviour research, in which process can be described only in unidimensional terms. Reliance upon LBDQ data collected from a single encounter with a group can clearly not provide adequate descriptions of leader behaviour and its relationships with group structure and process.

The general deficiency in leadership studies has been a lack of research designs that give information about group process changes over time. Classical controlled experimentation is not impossible in studies of leadership. Though the ease and convenience of *ex post facto* studies is readily acknowledged, we cannot overlook their very large weaknesses.[96] Greater attention, too, should be given to questions of methodology in explanatory surveys that are undertaken. In this regard, the techniques of multivariate analysis should be more generally applied as this concept is understood by Lazarsfeld.[97] The emphasis here is upon multivariate analysis in a structural sense where relationships between variables are examined and compared under different organizational settings or conditions. Statistical multivariate procedures are clearly relevant also, but are not identical. They give no assurance that partial or multiple correlations separate distinct types or groups of organizations so that varying organizational relationships may be compared within and between them. Similar attention should be given to wider problems of methodology and survey design.[98,99]

Conclusion

It is hoped that this review and critique will focus attention upon some issues in leadership research so that, while profiting from past research, investigators may consider issues often neglected or postponed as a result of the bloom of LBDQ-based research and psychometric survey designs. The paper stresses the importance of system-based concepts of organization, demonstates how a Canadian tradition of leader behaviour research has illumined some relationships in organizational behaviour, and suggests some theoretical and design problems to be considered in further research.

REFERENCES

[1] A revision of a paper originally delivered at the February, 1967 meeting of the American Educational Research Association in New York. The author is indebted to Miss Martha Schmidt, research assistant at OISE, for assistance in revision of this paper.

[2] R. M. Stogdill and A. E. Coons, (Eds.), *Leader Behavior, Its Description and*

Measurement, (Columbus: Bureau of Business Research, Ohio State University, 1957).

[3] R. T. Morris and M. Seeman, "The Problem of Leadership: An Interdisciplinary Approach", *American Journal of Sociology,* LVI, (1950).

[4] A. W. Halpin, *Manual for the Leader Behavior Description Questionnaire,* (Columbus: Bureau of Business Research, Ohio State University, 1957).

[5] A. W. Halpin and B. J. Winer, "A Factorial Study of the Leader Behavior Descriptions", in *Leader Behavior.*

[6] A. W. Halpin, *The Leadership Behavior of School Superintendents,* (Chicago: Midwest Administration Center, University of Chicago, 1959).

[7] J. H. M. Andrews, "Recent Research in Leadership", *Canadian Education,* XIII, No. 4, (September, 1958).

[8] L. von Bertalanffy, "General Systems Theory", in *General Systems: Yearbook for the Advancement of General Systems Theory,* L. von Bertalanffy and A. Rapaport, (Eds.), I, (1956).

[9] K. E. Boulding, "Some Questions on the Measurement and Evaluation of Organization", in *The Dilemna of Organizational Society,* H. M. Ruitenbeek, (Eds.), (New York: Dutton, 1963), pp. 133-134.

[10] C. A. Gibb, "Leadership", in *Handbook of Social Psychology,* G. Lindzey, (Ed.), (Reading, Mass.: Addison-Wesley, 1954), 11, 883.

[11] D. Cartwright, "Influence, Leadership, Control", in *Handbook of Organizations,* J. G. March, (Ed.), (Chicago: Rand, McNally, 1965), p. 39.

[12] J. K. Hemphill and A. E. Coons, "Development of the Leader Behavior Description Questionnaire", in *Leader Behavior: Its Description and Measurement,* R. M. Stogdill and A. E. Coons, (Eds.), (Columbus Bureau of Business Research, Ohio State University, 1957), p. 9.

[13] J. DeLamater, C. G. McClintock, and G. Becker, "Conceptual Orientations of Contemporary Small Group Theory", *Psychological Bulletin,* LXIV, (1965), p. 409.

[14] K. Yamamoto, "Analysis of Teaching: Another Look", paper read at American Educational Research Association, (Chicago, February 1966), pp. 9-10. See also a revision of this paper in *School Review, LXXV,* (1967), pp. 205-212.

[15] R. T. Morris and M. Seeman, *op. cit.*

[16] R. M. Stogdill, *Individual Behavior and Group Achievement,* New York: Oxford, 1959).

[17] L. Petrullo and B. M. Bass, (Eds.), *Leadership and Interpersonal Behavior,* (New York: Holt, Rinehart, 1961).

[18] A. Zaleznik and D. Moment, *The Dynamics of Interpersonal Behavior,* (New York: Wiley, 1964).

[19] M. B. Miles, "Planned Change and Organizational Health: Figure and Ground", in *Change Processes in the Public Schools,* R. O. Carlson (Ed.), *et al.,* (Eugene, Ore.: Center for the Advanced Study of Educational Administration, 1965).

[20] R. Dubin, "Supervision and Productivity: Empirical Findings and Theoretical Considerations", in *Leadership and Productivity,* R. Dubin, (Ed.), *et al.,* (San Francisco: Chandler, 1965).

[21] R. T. Morris and M. Seeman, *op. cit.*

[22] *Leader Behavior,* p. 1.

[23] C. A. Gibb, *op. cit.*

[24] D. Cratwright and A. Zander, (Ed.), *Group Dynamics,* Second edition, (New York: Harper and Row, 1960), pp. 489 ff.

[25] A. W. Halpin and B. J. Winer, *op. cit.*

[26] R. M. Stogdill, *Manual for the Leader Behavior Description Questionnaire: Form XII,* (Columbus: Bureau of Business Research, Ohio State University, 1963), p. 2.

[27] R. M. Stogdill, *Managers.*
[28] B. D. Anderson, "Leader Behavior Styles of Alberta School Principals", Unpublished Master's Thesis, (The University of Calgary, 1966).
[29] A. F. Brown, "Reactions to Leadership", *Educational Administration Quarterly,* III, No. 1, (Winter, 1967).
[30] K. F. Punch, "Bureaucratic Structure of Schools and Its Relationship to Leader Behavior", Unpublished Paper, (Toronto: Ontario Institute for Studies in Education, 1967).
[31] A. W. Halpin, *The Leadership Behavior.*
[32] *Group Dynamics,* pp. 492 ff.
[33] J. H. M. Andrews, *op. cit.*
[34] J. K. Hemphill, *Group Dimensions: A Manual for Their Measurement,* (Columbus: Bureau of Business Research, Ohio State University, 1956).
[35] *Ibid.*
[36] D. V. Morris, "Staff Characteristics and Principal Leadership", Unpublished Master's Thesis, (The University of Alberta, 1963).
[37] H. Thompson, "Leadership and the Provincial Superintendent", Unpublished Master's Thesis, (The University of Alberta, 1964).
[38] A. W. Halpin, *The Leadership Behavior.*
[39] A. F. Brown, *op. cit.*
[40] K. F. Punch, *op. cit.*
[41] R. M. Stogdill, *Manual,* p. 2.
[42] R. M. Stogdill, *Managers,* p. 27.
[43] A. F. Brown, *op. cit.*
[44] K. F. Punch, *op. cit.*
[45] A. F. Brown, *op. cit.*
[46] A. G. McBeath, "Teacher Leader Behaviour and Teaching Effectiveness", *Alberta Journal of Educational Research,* VI, (March, 1960).
[47] J. K. Hemphill and A. E. Coons, *op. cit.,* p. 35.
[48] B. T. Keeler and J. H. M. Andrews, "Leader Behaviour of Principals, Staff Morale and Productivity", *Alberta Journal of Educational Research,* IX, (September, 1963).
[49] R. M. Stogdill, *Managers,* p. 52.
[50] T. B. Greenfield and J. H. M. Andrews, "Teacher Leader Behaviour as Related to Student Growth", *Alberta Journal of Educational Research,* VII, (June, 1961).
[51] E. Miklos, "Dimension of Conflicting Expectations and the Leader Behavior of Principals", Unpublished Doctoral Dissertation, (The University of Alberta, 1963).
[52] E. Miklos, "Role Theory in Administration", *Canadian Administrator,* III, (November, 1963).
[53] M. Seeman, "Role Conflict and Ambivalence in Leadership", *American Sociological Review,* XVIII, (August, 1953).
[54] E. Miklos, "Dimension . . .", p. 1766.
[55] *Ibid.,* pp. 179-180.
[56] R. G. Fast, "Leader Behavior of Principals as It Relates to Teacher Satisfaction". Unpublished Master's Thesis, (The University of Alberta, 1964).
[57] B. D. Anderson and A. F. Brown, "Who's A Good Principal?", *Canadian Administrator,* VI, (December, 1966).
[58] *Ibid.,* p. 12.
[59] J. F. Punch, *op. cit.*
[60] W. G. Schmidt, "The Relationship Between Certain Aspects of Leader Behaviour and Organizational Climate", Unpublished Paper, (The University of Alberta, 1965).

[61] See J. H. M. Andrews, "Some Validity Studies of the OCDQ", *Canadian Education and Research Digest,* V, (December, 1965) for discussion of the Schmidt study.
[62] A. W. Halpin and D. B. Croft, *The Organizational Climate of Schools,* (Chicago: Midwest Administration Center, University of Chicago, 1963).
[63] A. F. Brown, *op. cit.*
[64] K. F. Punch, *op. cit.*
[65] H. Thompson, *op. cit.*
[66] K. F. Punch, *op. cit.*
[67] W. W. Charters, Jr., *Teacher Perception of Administrator Behavior,* Final Report, Cooperative Research Project No. 929, (United States Office of Education, 1964), p. 5.
[68] A. F. Brown, *op. cit.,* p. 72.
[69] J. H. Blocksidge, "Change of Leader Behaviour Attributable to a Leadership Course for School Principals", Unpublished Master's Thesis, (The University of Alberta, 1964).
[70] H. Thompson, *op. cit.*
[71] K. F. Punch, *op. cit.*
[72] D. A. MacKay, "Should Schools be Bureaucratic?", *Canadian Administrator,* IV, (November, 1964).
[73] R. Dubin, *op. cit.*
[74] A. Zaleznik and D. Moment, *op. cit.,* pp. 421-422.
[75] C. A. Bigg, *op. cit.,* p. 882.
[76] H. Hyman, *Survey Design and Analysis,* (Glencoe: Free Press, 1955), p. 178.
[77] D. T. Campbell and J. C. Stanley, "Experimental and Quasi-Experimental Designs for Research", in *Handbook of Research on Teaching,* N. L. Gage, (Ed.), (Chicago: Rand McNally, 1963).
[79] R. M. Stogdill, *Individual Behavior.*
[79] R. M. Stogdill, *Individual Behavior.*
[80] *Group Dynamics.*
[81] J. K. Hemphill and A. E. Coons, *op. cit.,* p. 35.
[82] G. C. Homans, *The Human Group,* (New York: Harcourt Brace, 1950), pp. 415-440.
[83] R. Likert, *New Patterns of Management,* (New York: McGraw-Hill, 1961).
[84] R. L. Kahn and D. Katz, "Leadership Practices in Relation to Productivity and Morale", in *Group Dynamics.*
[85] A. Zaleznik and D. Moment, *op. cit.*
[86] V. D. McNamara and F. Enns, "Directive Leadership and Staff Acceptance of the Principal", *Canadian Administrator,* VI, (November, 1966).
[87] M. B. Miles, *op. cit.,* p. 12.
[88] K. Yamamoto, *op. cit.*
[89] E. Amidon and A. Simon, "Teacher-Pupil Interaction", *Review of Educational Research,* XXXV, (1965).
[90] D. Cartwright, *op. cit.,* pp. 11 ff.
[91] A. Zaleznik and D. Moment, *op. cit.,* p. 432.
[92] R. M. Stogdill, *Managers.*
[93] A. Zaleznik and D. Moment, *op. cit.,* pp. 422-424.
[94] B. D. Anderson and A. F. Brown, *op. cit.*
[95] R. M. Stogdill, *Individual Behavior,* p. 33.
[96] F. H. Kerlinger, *Foundations of Behavioral Research,* (New York: Holt, Rinehart and Winston, 1965), pp. 290, 359.
[97] P. F. Lazarsfeld and M. Rosenberg, (Eds.), *The Language of Social Research,* (Glencoe: Free Press, 1962), p. 115.
[98] P. F. Lazarsfeld, "Problems in Methodology", in *Sociology Today,* R. K. Merton, L. Broom, L. S. Cottrell, Jr. (Eds.), (New York: Harper and Row, 1959), Vol. I.
[99] H. Hyman, *op. cit.*

BOOK REVIEWS

SCHOOLS OF THE FOOTHILLS PROVINCE: THE STORY OF PUBLIC EDUCATION IN ALBERTA

By: John W. Chalmers, Published for The Alberta Teachers' Association, (Toronto: University of Toronto Press, 1967).

This, the first history of Alberta's public school system, is the contribution of The Alberta Teachers' Association (ATA) to Canada's centenary celebrations. With the publication in 1964 of a similar history for British Columbia by F. H. Johnson, half of the public school story for western Canada has been told.

Unlike Johnson, Chalmers uses both the chronological and topical approaches, as did C. B. Phillips in his pioneer history of Canadian education a decade ago. Thus Part I divides the last one hundred years into ten periods (giving each a chapter), from the establishment of the first schools in the west in the early nineteenth century to the report of Alberta's royal commission on education in 1959. It is by far the most interesting part of the book. Part II, the curriculum section, traces the chronological development of several topics: high schools, vocational education, examinations and promotions, correspondence schools, special education, and Indian education. Part III, the administration section, does the same for centralization, bus transportation, teacherages, separate schools (Catholic and Hutterite), and educational finance. Part IV discusses the politics of education, reviewing, in turn, the evolution of the superintendency, the Alberta School Trustees' Association (ASTA), the Faculty of Education, and the ATA itself. An excellent index concludes the work, but there is no bibliography, although footnotes noting sources are frequent.

There is no doubt that the work is an invaluable contribution to the sparse literature on the development of education in western Canada. The ATA is certainly to be commended for having subsidized such a comprehensive study of the development of public education in Alberta. It will be a ready reference for many aspects of the subject in the province.

The work, however, is not without its serious weaknesses. Although presented as a "story", it is on the whole a dull story; and however much it may please the professional, it will not have a wide appeal. This is unfortunate for the intention, as shown in the choice of terminology ("story" rather than "history"), was undoubtedly otherwise. One cannot fault the style, for the book is well-written. It may be tedious but it is not difficult to read. The work suffers, however, because it is primarily

the story of institutional growth—the growth of the public school system rather than public education, which, as Bernard Bailyn (*Education in the Forming of American Society,* 1960) has made clear, is a much wider term.

In the institutional approach personalities are unimportant, or are at least of minor significance, and events take place as if men did not exist. Instead of correspondence (private or official) and diaries or memoirs (where these exist), or even newspaper editorials and letters to the editor, the main sources of information in the institutional approach are the colorless and tiresome official government publications, committee reports, and university theses. Yet in the work under review the story would be much brightened if men such as Perren Baker, H. C. Newland, J. W. Barnett, G. P. Smith, T. O. King, G. F. McNally, and Wm. Aberhart —all prominent in the story—were delineated as people rather than presented as incidental ciphers in the course of Alberta's educational development. In new regions such as Alberta contemporary histories are probably inevitable, but their blandness (to protect the living) might be partially overcome by consulting the same living about the philosophies, aspirations, and motivations of the late dead. This is particularly important where the few sources of important correspondence that are known to exist (the files of the ATA, the ASTA, the Faculty of Education) have apparently not been consulted.

There is perhaps another reason why the work is not as interesting as the author's well-known wit and not inconsiderable talents as a writer would lead one to expect. The work pays little attention to the feelings, hopes, and thoughts of the people who settled Alberta. Thus the people came in the main from Ontario, the British Isles, the United States, and central and eastern Europe, yet there is no treatment of the ever-present problem of Canadianization and the resulting interplay of Ontario, British, and American influences on the work of the schools. The Ontario heritage with its WASP overtones is largely overlooked, as is the impact of such mid-west phenomena as populism, religious fundamentalism, progressivism, and the neo-conservative reaction of recent years. It is doubtful whether educational history will ever amount to much until it becomes a handmaiden of social history—until it takes the ideals of the people and the motivations of their leaders in the establishment of institutions into account. The work under review fails in this respect. It fails to see education in its "elaborate, intricate involvements with the rest of society" and to relate these to education's "shifting functions, meanings and purposes" (Bailyn, *op. cit.,* p. 14). As a result, the first cultural history of education has still to be written in Canada.

M. R. Lupul
The University of Alberta

LEARNING TO READ: THE GREAT DEBATE

By: Jeanne Chall, (New York: McGraw-Hill Book Company, 1967).

Sponsored by the Carnegie Corporation, Dr. Chall initiated in 1962 and completed by 1965 a critical analysis of existing research "comparing different approaches to beginning reading". Her general objective was to examine current innovations and determine whether ". . . these changes . . . [are] justified by existing evidence or by the results of current experimentation". More specifically, she hoped that the study would also ". . . salvage whatever we already knew, . . . help point up specific gaps in our knowledge . . . and help future experimenters design more crucial, meaningful studies".

Publications by Nila B. Smith (1934, 1960), Fries (1963) and Barton and Wilder (1964), as well as by others equally eminent in their respective fields, had already gained professional acceptance for the conclusion that the vast, fragmented accumulation of reading research is inadequate as a foundation for future research. Chall's 371 page discussion of the "rather dismal picture" which she says research in reading presents, would have to be considered a redundancy unless it had something to offer that would take it beyond the writings of these others. When, in addition, one stops to consider the fact that practice in the field already incorporates many of the "innovations" which Chall suggests to improve materials, methods and research, one might also be tempted to conclude that, from the point of view of reading research, this book was obsolete before it was published.

Dr. Chall explains, however, that *Learning to Read* is a revision of her study, and that it appears in its present form to make it "useful to educators and understandable to laymen". It is, therefore, from this point of view that the book must initially be evaluated.

What has it to offer "educators" (inclusive as such a term may be)?

The book does present a sequential description of the historic controversy over method that seems to have become chronically associated with the teaching of beginning reading. As might be expected from the co-author of one of the most respected formulas for predicting readability (Dale-Chall), the book, despite being closely packed with detail, is clearly written. It is free of professional jargon, it has been prepared with great care and professional integrity and its documentation is excellent. Terminology is carefully defined. Included are italicized pre and post summaries for each section in the form of questions which lead to more questions and summaries at the conclusion of each major topic, as well as generalizations drawn, as possible hypotheses for future research, at the end of each major section.

The presentation of the pros and cons of the debate is done with great subtlety: partisan discussion has been predicted and anticipated by the provision of answers, in the body of the text, to almost any question or argument that might be advanced. The frequent use of tentative state-

ments is deceptively persuasive in leading the biased reader to withhold judgment. This sophisticated technique would seem to be both necessary and effective in the discussion of controversial issues, because as Chall says:

> Since neither the issue nor the evidence was clear-cut, I describe, whenever possible, the process of reasoning I went through to arrive at my interpretations.

The book also provides criteria and schedules by which an educator can evaluate research, innovations, basal readers and other materials, lessons, activities or classroom procedures generally. This section, particularly, is excellent.

The author's conclusions about innovation and change and the pressures which result in the implementation of or resistance to these processes, from the educator's viewpoint, are also discussed at some length. As a result the educator is, perhaps, enabled to see himself in perspective, in the context of his professional relationships with school, parents, community, textbook publishers and teacher training institutions. He can also gain insight into his failings and possible failings as an innovator and interpreter of reading research. Thus, in addition to the valuable anaylsis and "essencing" of past research, the author has provided a reference that is enhanced by the inclusion of current laboratory studies as well as the most recent large scale projects on beginning reading.

The educator, therefore, will find this book of great value, and in this area the author can be considered to have been eminently successful in fulfilling her stated objectives.

What of her intentions regarding the layman?

While lacking a professional orientation toward the question, the layman, too, could perhaps garner considerable information from the author's presentation. The description of innovations, the comparison between British and American teacher attitudes to children or theory, and the discussion of basal readers is topical and pertinent. One might, however, wish to be assured that the layman to whom this material is made so freely available would be able to assess, interpret and use it as objectively as one relies on a professional to do. The fact that he is expected to benefit from the "understandable" italicized summaries at the beginning and end of each chapter does not in any way deter him from misinterpreting or supporting from the reported research biases which hitherto had rested on opinion alone. The advisability of directing professional publications of this nature co-jointly to such widely disparate audiences can therefore be seriously questioned.

Reading Dr. Chall's book was an interesting and enjoyable experience and yet it resulted in an aftermath of disquietude. Reading specialists are still analyzing their field "parochially" rather than conceptually! Despite the fact that she realizes that:

> . . . more research may engender more confusion if it is not synthesized and put into a meaningful framework . . .

and that

> . . . when new reading programmes are adopted they should be cast in a research and development framework . . .

Chall herself fails to go beyond conventional limits. She does succeed in consolidating under two major headings ("Code Emphasis" and "Meaning Emphasis") the splinter groups which constitute the ranks of methodological theoreticians. Had she, herself, with her background of experience, knowledge and professional status, however, seen fit to develop a conceptual framework for analyzing and synthesizing her data, rather than restricting herself to the usual methodological terminology, *Learning to Read* might have been a major contribution to educational literature.

D. M. LAMPARD
The University of Lethbridge

THE PSYCHOLOGY OF LANGUAGE, THOUGHT, AND INSTRUCTION

By: John P. De Cecco, (New York: Holt, Rinehart and Winston, Inc., 1967).

The editor of this excellent book of readings has been successful in drawing together articles, representing a wide variety of related disciplines, which will further the general objective—greater knowledge and understanding of instruction in language and thought.

Chapter 1 deals with the interdisciplinary field of psycholinguistics: a consideration of the communicator as he encodes and decodes messages. Here transformational grammar is included because Chomsky's descriptions are based on the psychological processes which generate sentences. Other grammars are supposedly not included because these are *after the fact* descriptions of language products—products which might, nevertheless, communicate the results of linguistic and cognitive activity. While not arguing with the editor's rationale in this case, it seems likely that other descriptions of language, such as that of the structural linguist, could have been used to suggest that children at an early age are capable of using language which must have been accompanied by higher mental processes.

In Chapter 2 the editor has assembled a theoretical base for a consideration of the way in which language systems control thought. Such a consideration of linguistic relativity might have been strengthened by the inclusion of some of the available research evidence in this area and by linking it to Chapter 3, the relation of language and thought to social class differences. It would seem likely, in this regard, that some of the

careful investigations of the language of the culturally disadvantaged would support the somewhat less secure claims of the linguistic relativists. At this point the reader must sympathize with the editor: so little is available to provide the synthesis he tried for.

Chapters 4, 6, 7, and 9 represent psychology. These deal with the nature of meaning in language and with ways in which meaning at the conceptual and propositional levels is acquired. This section forms the solid core of the book. De Cecco has chosen his readings carefully to represent fairly both the cognitive and behaviourist schools of psychology. He has, perhaps, done all that might be done at this point in time to present an overview of the psychology of language.

Chapter 5 is one of the two chapters dealing with instruction and is the first real disappointment in the book. The editor might certainly be excused for choosing reading for consideration with its extensive research and theoretical literature. It is, however, difficult to justify a view of reading that includes little but word calling and this in a book with a psychological emphasis. Perhaps the real error in judgement was in not putting reading in its proper place within the language arts. A consideration of the encoding which might reasonably accompany the reading act would also have ensured the inclusion of higher mental processes and, in spite of the definition of reading used, provided a more worthy example of instruction in language and thought.

The second chapter dealing with instruction, Chapter 10, lacks something, too, in the satisfaction it provides the reader, but here the editor is not to be faulted. The authors themselves have seached for ways of providing instruction in language and thought and, while their attempts are often exciting, the reader is left with the feeling that there is a long way to go both in building instructional models in this area and in verifying them.

This is an interesting and a useful book. It is much more than a collection of readings which, on the whole, are well selected. The editor's excellent introductions to each of the chapters provide a theoretical framework for the chapter itself and for the book. Such was the inadequacy of the literature available for the editor's consideration in the case of some of the chapters, however, that a reader might often wish for even more in the way of a synthesis, pointing wherever possible to next steps in theory and research.

R. D. Armstrong
The University of Alberta

GREEK ETHICS

By: Pamela Huby, (London: Macmillan & Co. Ltd., 1967).

ETHICAL INTUITIONISM

By: W. D. Hudson, (London: Macmillan & Co. Ltd., 1967).

CONTEMPORARY MORAL PHILOSOPHY

By: G. J. Warnock, (London: Macmillan & Co. Ltd., 1967).

EXISTENTIALIST ETHICS

By: Mary Warnock, (London: Macmillan & Co. Ltd., 1967).

These four books are the first to appear in a new series, *New Studies in Ethics,* edited by W. D. Hudson. The series as a whole is intended to "meet the need for a comprehensive survey of ethics from the point of view of contemporary analytical philosophy" (editor's preface to the first book). On the evidence of these four books, the series promises to be an interesting and important one, both from the point of view of the professional student of philosophy and the interested layman. I am unclear as to the purpose of the editor's emphasis on "contemporary analytical philosophy" as the point of reference. It is true that the contributors to the series are all well-known philosophers associated with that tradition, and thus the categories employed tend not to be those of German or French philosophy. However, I find the books reviewed mostly free of ideological axe-grinding, and written in language that requires no particular familiarity with technical philosophy. I propose to take each book individually and then make some general comments on them as a group.

Mrs. Huby offers us a survey of Greek ethical thought from the beginnings through to the Stoics, the last Greek school of philosophy. She treads very skillfully the narrow ridge between unwieldy scholarship and unhelpful popularizing: the more esoteric material is relegated to footnotes at the end of the book. Certain points which are crucial to the understanding of Greek ethics are made very clearly, and the resulting picture is faithful to the documentary evidence without being enslaved by it. She makes five divisions, pre-Socratic thought, Socrates, Plato, Aristotle and post-Aristotelian, and in this is reflected perhaps the one item where she shows impatience with the rigours of scholarship. The chapter on Socrates contains many overt speculations as to the extent and content of his philosophy; while the ideas she attributes to him are interesting and valuable, the evidence for them I find too scanty. The evidence suggests Socrates was important for his method rather than for the content of his philosophizing; Mrs. Huby tries to give him both.

In dealing with earlier Greek thought, she emphasizes well the role of history in its development. Greek morality puts a high value on

results not motives as being of ultimate moral significance, and on individual happiness as the end of life. These are notions abhorrent to people brainwashed by the nobility of suffering and "It matters not who won or lost, but how you played the game". This leads as much to unfair and hysterical evaluations of them as immoral as it does to Nietzschean trumpet-calls for the return to the healthy ideas of our Greek forefathers. The truth is that the appearance of this value-scheme at that stage of Greek history was as determined by the nature of their society, as ours is to a large extent by our society. Physical survival of Greek communities for a number of decades was not possible except under that scheme of values, and comparisons made in ignorance of the fact, whether favourable or unfavourable, are inappropriate.

Plato is well presented as inspired by but yet surpassing Socrates, and as challenging the accepted morality in a fundamental way. Mrs. Huby accuses him of too much internalizing virtue, of assuming without proof that, if intentions and motives are right, right acts will follow (p. 31). Plato does claim this, but he offers a justification, namely that Virtue is Knowledge and Knowledge is absolute and infallible. Against this metaphysical background, the internalizing of virtue is not in itself illegitimate. However, the metaphysical background is unacceptable, and from this it follows that the internalizing is unacceptable.

In two places, the relevance of the Greeks to contemporary moral theory is brought out. Mrs. Huby spends sometime on the *Philebus,* the last real dialogue Plato wrote. It is a philosophical analysis of pleasure which is worthy of consideration by anyone interested in Hedonism, and is well presented in this book. Aristotle also deals with some of the issues being presently debated, whether and, if so, how there is an essence of goodness, the importance of the particular situation in ethical decision-making, the logic of responsibility, the relation between thought and action and, again, pleasure as an ethical value. Not only does Aristotle debate these issues, he also, as Mrs. Huby indicates (p. 44), debates them in a manner analogous to that of many contemporary philosophers, the manner epitomized by John Austin's remark that "ordinary language is *not* the *last* word. It *is* the *first* word" (*Philosophical Papers,* p. 133).

W. D. Hudson's book is, I think, the best of the four. By "Intuitionists" he does not mean the 20th century group led by Moore and Ross, but the group of 18th century British philosophers consciously reacting to Hobbes, namely Shaftesbury, Hutcheson, Cudworth and the rest. Hudson is content to let these men speak for themselves, and by careful arrangements of their thoughts, he is able to make this tactic totally successful. For what is remarkable about this group of philosophers is their apprehension of the strengths and weaknesses of their own position. While they tend to agree on the strengths, they attempt to meet the weaknesses in different ways. Although, as Hudson says in his concluding assessment of the picture thus far presented, there are funda-

mental objections to any form of Intuitionism, the liveliness of this private debate among a group sharing common presuppositions is quite fascinating, and shows that Intitutionism need not necessarily be the kind of sterile philosophy it is often thought that is must be.

The central point at issue here is one that is central in Plato and Aristotle too, namely the relation of thought and action. How can an intuition, which is primarily cognitive, issue in moral action, which is primarily a matter of will and not just reason? Many argued that the apprehension of the end to be sought moved the will, whetherf this apprehension was a matter of the senses or of the reason. None of them was prepared to accept the suggestion that many contemporary theorists put forward, that the will determines the reason. Butler compromised by making Conscience combine the roles of reason and will. However, there is the obvious objection here, and I cannot see that Hudson makes it in so many words, that this theory simply defines away the problem in a manner no one will find satisfactory.

Hudson also makes a point with respect to these philosophers analogous to one made by Mrs. Huby about the Greeks, that their views are a product of their social context. He points out that they were writing at a time when there was "a high degree of unanimity in the deliverances of conscience" (p. 61). Thus the empirical facts seemed to support an Intuitionist theory, in the same way as the empirical facts of morality do not now. We are now, for various reasons which are not relevant here, aware in a way we were not two hundred years ago of the differences in moral beliefs of different people and different ages. This is widely held to form the same sort of support for a subjectivist, whether personal or social, account of morality as the unanimity was held to form for Intuitionism. What needs to be emphasized here, and Hudson does not mention it, is that either opinion is confused. The thesis that moral knowledge is a matter of the intuition of objective values, or of the subjective creation of values are philosophical and so *a priori* theses. Facts about agreement or disagreement among men with respect to moral matters are as irrelevant to the truth of such theses as *any* facts to the truth of *any a priori* thesis.

The question of the sterility of Intuitionism will serve as a bridge to Mr. Warnock's book, for it is here that Moore *et al.* are mentioned. I will not say "discussed", for Warnock proves to his own mind the sterility of this brand of Intuitionism by spending little time on its propounders. It is true that a poorly articulated defence of Intuitionism will be sterile, but Warnock is too ready, to my mind, to find in Moore and Ross such poor articulation. It is also true that Intuitionism is easy to ridicule: few instructors in ethics, and the present writer is not one of them, have failed so to play to the gallery. But what is easy to ridicule is not necessarily ridiculous. Warnock seems emotionally committed to denigrating this movement, remarking quite gratuitously that "Ross deviates in the direction of good sense" (p. 11), and that Moore

is guilty of "a rather unpleasing and even arrogant self-assurance" (p. 5); I find Warnock manifestly guilty of this fault himself. He culpably misrepresents Moore at one point, in approving of Ross' view that an object's value is supervenient on its other properties (one of Ross' putative "deviations") and implying Moore was too stupid to see this point in his time. That is just false. Moore goes into the whole issue in some detail in his paper "The Conception of Intrinsic Value", published in 1922, some eight years before Ross' views appeared in print; this hardly justifies attributing this improvement on Moore's original doctrine to Ross alone. In short, the fact that contemporary moral philosophy in the analytic tradition began with Moore, and that he had and still has an enormous influence, are both represented by Warnock as being entirely regrettable, and this is a gross injustice to a man who said what he said because he could see no other defensible alternative. What Moore says at times seems odd, but it is his shrewd understanding of his subject-matter which led him to say it that is his real contribution, and this Warnock effectively ignores.

Other movements and figures get similar short shrift. Emotivism is summarily dismissed, and once again Warnock erroneously implies that C. L. Stevenson is quite unaware of certain criticisms of his theory, e.g., that on p. 27 about some uses of moral language being not in fact emotional. In *Ethics and Language* this and other of Warnock's points are anticipated and dealt with by Stevenson. Hare's prescriptivism is given a longer run, but the objections are old hat, and many moral philosophers have got beyond dealing with Hare at that level. There are also some surprising omissions. The book is supposed to "provide a compendious survey of moral philosophy written in English since about the beginning of this century" (p. 1), and yet there is no mention either of Dewey's Instrumentalist theory, or of the cultural and individual subjectivism of Westermarck and Perry. It is true that these thinkers have not had much influence in Britain, but they had and still have a great deal of influence in North America. Perhaps Warnock does not think English is spoken in North America. Toulmin, who, like Hare, sought to build on Intuitionism and Emotivism, is barely mentioned; yet he followed a different and in may respects more interesting path than did Hare from the same starting point.

The discussion of these three viewpoints occupies slightly more than half the book. For the remainder we are treated to Warnock's opinions on what contemporary moral philosophy ought to have concerned itself with, rather than what it did concern itself with. What there is of significance in the book is in this part, even though it is, as Warnock concedes, somewhat disorganized. He makes some helpful remarks (Chap. VI) on the opposition to Naturalism, which corrects some misunderstandings about that debate. Overall, however, this seems to me the least satisfactory volume of the series. An extended sympathetic report on the tradition would have achieved at least as much as War-

nock's unsympathetic cursory review and his own gospel, would stand much less risk of alienating readers, and would be more in keeping with the purpose of the series.

Mrs. Warnock attempts to be quite fair in her account of Existentialist ethics; she does not conceal her fundamental disagreement with the type of philosophizing she is discussing, but neither patronizes them as sincere but naive, nor inveighs against them as malicious and disruptive. She acknowledges the enormous difficulty in presenting a brief account of any Existentialist thought; its deliberate method tends to be unsystematic and to emphasize the individuality of entities and contexts. Any summary must necessarily ignore this, and lose much of what is valuable. However, as Mrs. Warnock rightly maintains, there are some basic presuppositions in Existentialist ethical theory, and she sets herself to isolating these together with a few examples of these presuppositions at work in particular cases.

Two main themes are traced out; firstly, the connection between ethics and the rest of philosophy in their thought, and secondly the ramifications of the claimed uniqueness of every moral situation. Clearly there is a close connection between their ethics and their general philosophy: the uniqueness of the individual is such that in all respects *the* world is *his* world. Mrs. Warnock sees the connection as flowing from general philosophy to ethics—"All philosophy leads ultimately to an answer to the ethical question" (p. 53). There is one slight danger in bringing out this point, which is to forget that the general epistemological basis from which Existentialist thought derives is significant in its own right, not merely because of the ethics to which it leads, and, for Kierkegaard and Heidegger at least, is seen as significant and leading to conclusions in non-ethical fields of enquiry.

Mrs. Warnock spotlights unerringly the difficulties I find with Existentialist ethics. What is emphasized is the solitude of the individual, the uniqueness of each decision that individual is called to make; the result is that little can be done in the way of specifying the content of morality except to say, Be authentic, Be yourself. This, she says (p. 54-55) involves making solely the act of moral choice ethically significant, without regard for what is chosen. This is absurd, for "if choosing freely for oneself is the highest value, the free choice to wear red socks is as valuable as the free choice to murder one's father" (p. 54). The second difficulty is closely related to the first. From the same viewpoint is derived the thesis of absolute responsibility—we are always responsible for our choices, no matter what the context of our choices. This is, interestingly enough, a thesis found in Aristotle, though arrived at from a somewhat different angle. This is also absurd, obliterating distinctions that need to be made. To glory in the absurd does not diminish its absurdity.

The one single thing that emerges most powerfully from these four books, dealing as they do with theorists of widely different ages and

contexts, is how much even so are there common concerns and common viewpoints. There is a ready explanation for this: ethical philosophy is an enquiry whose area is man and not his world, man as having dimensions to his nature not to be captured in impersonal mechanical laws. The progress of science and technology, the supplanting of old knowledge by new, are in this respect different, and both types of enquiry should, for their mutual benefit, not seek to concern themselves with the other's business. If we are still chewing the same things over as did Plato and Aristotle, this might imply that we should give philosophy up. To agree with that view would be to make the mistake I have just outlined, to assume erroneously that there is something wrong if we have not moved on to other problems. Philosophical problems, in whatever field, education included, do not go away, and progress in philosophy is not a matter of making them go away. It is a matter of understanding those problems and why they have on us the perennial grip they do have. In one area of philosopsy, this new series looks like teaching that lesson well.

Roger A. Shiner
The University of Alberta

NOTES ON CONTRIBUTORS

SANTOHK S. ANANT holds undergraduate degrees and a master's degree from Panjab University and a Ph.D. from the University of Michigan. He is an Associate Professor of Psychology at the University of Lethbridge. He has published several papers and contributed to a book entitled *Environmental Abstracts,* A University of Michigan publication.

C. C. ANDERSON, a graduate of the Universities of Glasgow and Cambridge, is a professor of Educational Psychology at The University of Alberta. In 1965-66 he was a Visiting Fellow in Social Psychology at Princeton University and Visiting Professor at Educational Testing Service. He has published papers in *Behavioral Science, Journal of Abnormal and Social Psychology, Journal of Personality and Social Psychology, Journal of Genetic Psychology, British Journal of Psychology* and is a co-author of a book of readings, *The Cognitive Process.*

R. D. ARMSTRONG holds bachelor's and master's degrees from the University of Alberta and an Ed.D. from the University of California at Berkeley. He was formerly a teacher and principal in elementary schools of Alberta and is presently an Associate Professor in Elementary Education at the University of Alberta. He is the senior editor of a monograph on Teacher Education, a report of the invitational conference sponsored by the Department of Elementary Education, which is to be published soon.

T. B. GREENFIELD, is Chairman of the Department of Educational Administration at the Ontario Institute for Studies in Education. Prior to his position as chairman he was Associate Professor in the same department. He was previously Research Director for the Canadian Teachers' Federation. He holds a B.A. from the University of British Columbia and M.Ed. and Ph.D. degrees from The University of Alberta.

DOROTHY M. LAMPARD holds a B.A. degree from the University of Manitoba, an M.A. from the University of Chicago, and an Ed.D. from the University of Oregon. She was a member of the Faculty of the Department of Elementary Education at The University of Alberta before becoming Professor of Education at the University of Lethbridge. Her special interests are learning and behavioural problems of children, and teacher education.

THOMAS E. LINTON is a Professor of Education and a Director of Disturbed Children's Program, University of Wisconsin, Milwaukee. He collaborated on a book of readings for the Social Foundations of Education field and is a co-author of *Patterns of Power: Social Foundations of Education.*

M. R. LUPUL is an Associate Professor in Educational Psychology at The University of Alberta. He has written on social studies and education in Western Canada in the *ATA Magazine, Canadian Educational and Research Digest,* and in *Comparative Education.* His latest publication, *The Campaign for a French-Catholic School Inspector in the Old North-West Territories,* 1893-10, was in the most recent issue of the *Canadian Historical Review*

R. S. MACARTHUR, whose doctoral study was carried out at the University of London, has for some years been Professor of Educational Psychology at The University of Alberta. His current research interests are on the theme of cognitive development in cross-cultural context. His numerous publications include reports of work with native peoples of the Canadian North and of Central Africa.

L. D. NELSON obtained his Ph.D. from the University of Minnesota after graduating from The University of Alberta. He is co-editor of MATHEX, published by Encyclopedia Britannica of Canada. He is a member of the Publications Committee and a member of the Editorial Board for a yearbook on the history of mathematics, National Council Teachers of Mathematics.

DAIYO SAWADA holds a joint appointment as an Assistant Professor in Elementary Education and in the Division of Educational Research Services at The University of Alberta. He has had articles published in the *Mathematics Teacher.*

ROGER A. SHINER did undergraduate work at the University of Cambridge. He has M.A. degrees from The Universities of Alberta and Cambridge and is presently completing a Ph.D. in Moral Science for Cambridge. He is an Assistant Professor of Philosophy at The University of Alberta.

The Alberta Journal of Educational Research

Vol. XIV, No. 2 *June, 1968*

CONTENTS

BOOK REVIEWS

COMMITTEE ON EDUCATIONAL RESEARCH
The University of Alberta

A SPECIAL ISSUE
THE USE OF COMPUTERS IN EDUCATION

A special issue on The Use of Computers in Education is being planned for 1969. Several manuscripts are in hand but scholars interested in this area are invited to submit manuscripts or give notice of intention to do so.

Editorial Comment

The funding of research in Canada is a problem of major proportions. It is particularly difficult to get funds in sufficient quantity to research problems in education. New agencies, such as the Human Resources Research Council in Alberta, have been created to fund research into social problems, including education. Will these agencies be able to provide enough money to finance all of the research necessary in these areas? Perhaps more important, will the agencies be able to fund the projects that qualify as basic-research as well as the developmental projects?

The term "basic-research" is difficult to define. It has been variously described as "when you don't know what you're doing," as "the opposite of practical research," and as "farfetched and useless." It could be said that basic-research is a creative synthesis; it is not seeing but discovering. However, it is probably not as important to define basic-research as to be able to differentiate between basic-research projects of greater and lesser importance.

The greatest need in any field of research endeavour is to recognize significant basic-research in its earliest stages before any practical applicability is apparent. There is the need to convince those charged with dispending the funds that the research *is* important. Research based on deductive reasoning, more properly called development, is relatively easy to finance because of its immediate applicability to everyday problems which can be described precisely in a routine application for funds. Those involved in basic-research have the problem that original or creative research cannot be planned in advance. Because basic-research is discovery, it cannot be anticipated wholly on the basis of known facts and generalizations.

Researchers investigating fundamental problems in education must accept responsibility for translating their problems into a language meaningful to the layman. The layman, in turn, must realize that, however simplified, the essence of basic-research cannot be assimilated without effort on his part. Can educational researchers give assistance to funding agencies in establishing guides for determining which are the better basic-research projects? Selye[1] uses three criteria to identify significant scientific *discoveries.* Is it possible to restate these criteria so that they might form a useful basis for making decisions about projects in their early stages? Selye states his criteria in this way:

> . . . it is characteristic of great basic discoveries that they possess, to a high degree and simultaneously, three qualities: they are true not merely as facts but also in the way they are interpreted, they are generalizable and they are surprising in the light of what was known at the time of the discovery [p. 149].

The importance of the word *simultaneous* must be emphasized. Selye illustrates, from the field of medical research, discoveries that would

1 See an article by Hans Selye, a prominent medical researcher, in Thruelsen, R. T. and Kobler, J. (Eds.). *Adventures of the mind,* New York: Vintage Books, 1960.

satisfy one or two of the criteria, but not all three. These would not, in his opinion, be classified among the great discoveries.

It is not sufficient for researchers to complain about the lack of funds for educational research, particularly for projects that would be classified as basic-research. They must be ready to help in stating criteria by which funding agencies can identify the projects most worthy of support.

* * *

A special issue of the *Alberta Journal of Educational Research* is being planned for 1969. As in the issues devoted to models in education and to the work of Piaget, this special issue will be on one theme: The Use of Computers in Education. Several manuscripts on this theme are already on hand; others will also be considered for inclusion in the special issue.

Articles of two types seem pertinent to the special issue. One type would be broadly based, setting the stage for an insightful look at the total area of computers and education. The second type would include articles on the use of computers in the research process in education or in the instructional process, including evaluation. Interested scholars are invited to submit manuscripts related to this theme, or to give notice of their intention to do so.

* * *

In this issue of the *Journal,* Gupta and Hemphill explore the theoretical aspects of two diverse educational questions. Gupta discusses the appropriateness of traditional methods of item analysis and presents theoretical support for the use of multivariate techniques for analyzing item responses. Hemphill argues for the need to look at innovativeness within the larger framework of general systems theory.

MacArthur presents a synthesis of several projects investigating methods of assessing the intellectual potential of Canadian native pupils. He summarizes data obtained in a series of longitudinal studies carried out under his direction.

Andersen and Stiles report a study of the use of expectancy tables in advising students in small rural high schools. Counsellors will have to weigh the convenience of such an approach against its use by untrained personnel.

The article by Roden and Slakter reports an empirical study of the effect of risk taking on student scores on multiple-choice examinations. They compare the student's perception of the risk he takes, and his willingness to do so, with the statistical evidence of his tendency toward risk taking.

In the last issue of the *Journal,* Sawada and Nelson presented an argument for a method of assessing length conservation that was independent of verbal facility. In this issue they present the results of a study utilizing this method. Their results are compared with the results of other researchers using a verbally-dependent method to determine conservation.

P.A.L.

Alberta Journal of Educational Research
Vol. XIV, No. 2, June, 1968

The author presents empirical evidence derived from cluster and factor analyses of item responses in support of the argument that for developing internally consistent or homogeneous measuring instruments, multivariate techniques for analyzing responses should get precedence over the traditional methods of item analyses.

RAM K. GUPTA
The University of Alberta

Multivariate Analyses of Test Responses as a Pre-Requisite to Item Analysis

Problem

Item analysis of one type or the other is universally regarded as an indispensible step in developing objective tests. An important objective of item analysis is to ensure internal consistency (also called homogeneity and coefficient alpha) so that the test may consistently measure "whatever it purports to measure" or "whatever it really measures." The phrases within quotes refer respectively to the claimed and the actual *validities* of the test. Ideally, criterion measures should be obtained independent of the test being developed which may be intended to be a more efficient substitute for the procedure leading to the criterion measures. Since external criteria having acceptable reliability are rarely obtainable, recourse is generally taken to an internal criterion—the scores on the test itself, the items of which are being analyzed.

The present research poses two questions. The first relates to item analysis against criteria in general—external or internal—and can be stated as: "Is it posisble to develop a homogeneous test if item analysis is done, using a heterogeneous criterion?" Logically, the answer would seem to be in the negative.

The second question is specifically related to the use of the scores on

the test itself as an internal criterion which is by far the commonest criterion used for item analysis. Since the test scores are based on the responses to the individual items, accepting these scores as an internal criterion should logically imply that the individual items also are acceptable, in which case item analysis would be superfluous. On the other hand, a lack of confidence in the adequacy of the items would imply that the test scores also constitute an inadequate criterion against which item analysis would be neither desirable nor justifiable. Is there a way out of this dilemma or vicious circle?

The present research seeks an answer to the above questions simultaneously insofar as they relate to the use of internal criteria for item analysis.

The problem as posed above points to the desirability of making the criterion homogeneous before using it for item analysis. When the criterion is internal, its homogeneity can be examined easily through multivariate analyses of the responses to the items constituting the criterion.

Data and Procedures

The data of this research consisted of the responses given by 460 girls and 554 boys to the fifty items of the Sequential Tests of Educational Progress, Mathematics, Form 2A (STEP-Math. 2A). The subjects were students of ninth grade mathematics in twenty-nine schools in the U.S. mid-west.

The reliability of the test in terms of Kuder-Richardson formula twenty (K-R-20) was found to be .71 for girls and .75 for boys. In view of the facts that (*a*) the subjects were substantially heterogeneous, (*b*) the instrument was a widely used standardized test, and (*c*) the nature of the subject matter would normally lead to high reliabilities, the obtained figures seemed rather low. It was decided, therefore, to examine the test for homogeneity before the test scores were used for item analysis. Five different multivariate analyses were used for this purpose. Each was applied separately to the responses of the boys and those of the girls. They were:

1. The method of "homogeneous keying" of DuBois, Loevinger, and Gleser (1952), applied to the inter-item matrices of variances and covariances,
2. principal factor analysis of phi-coefficients,
3. principal factor analysis of tetrachorics,
4. alpha factor analysis of phi-coefficients, and
5. alpha factor analysis of tetrachorics.

In all the factor analyses, squared multiple correlations were used for communalities wherever possible, four factors were extracted, and the resulting common factor space was subjected to normal varimax rotation.

The rationale for selecting the above methods of multivariate analyses and a detailed comparison of the results given by them forms the subject matter of another paper now under preparation. It is sufficient to state here that each of them revealed that the test was undoubedly heterogeneous, containing two homogeneous subgroups of items representing two independent abilities, leaving outside their fold about a third of the fifty items.

The above was followed by steps 1 through 4 given below:

1. Using the scores on *all the items* of the test as the internal criterion, biserial correlations were calculated for each of the fifty items;

2. Using the scores obtained *from the items falling in a given factor* as the internal criterion, biserial correlations were found for only those items which were included in the factor;

3. Calculations corresponding to 2 above were done using scores on each of the clusters as the internal criterion;

4. Calculations of K-R-20 reliabilities were done for the results of analyses 2 and 3.

Results and Conclusions

Table 1 shows the biserial correlations and K-R-20 reliabilities for boys' responses when all the items constituted a single internal criterion and also when the biserial correlation of an item was calculated with reference to that segment of the internal criterion for which the item had shown maximum affinity, considering the results of multivariate analyses. Table 2 contains the corresponding information for girls.

The results of Tables 1 and 2 led to the following conclusions:

1. Factor analyses split the test into three parts containing 17, 19, and 14 items in the case of boys and 15, 18, and 17 in that of girls. Similarly, the method of "homogeneous keying" split the test into three parts, containing 14, 17, and 19 items in the case of boys and 15, 18, and 17 in that of girls (Vide second last row in Tables 1 and 2). The first two factors and clusters were highly homogeneous; the third was extremely heterogeneous, as evidenced from the reliabilities given in the last row of the tables. In fact, the heterogeneous part merely contained the items not included in the homogeneous clusters or factors.

TABLE 1

BISERIAL CORRELATIONS AND K-R-20 RELIABILITIES FOR BOYS' RESPONSES TO STEP-MATH 2A CLASSIFIED BY NATURE OF CRITERION SCORES

		Nature of Criterion Scores					
Item no.	Total test	Fact. I	Fact. II	Fact. III	Clus. I	Clus. II	Clus. III
1	—.053			.332			.192
2	.363	.510					.358
3	.500	.715			.720		

Item no.	Total test	Fact. I	Fact. II	Fact. III	Clus. I	Clus. II	Clus. III
4	.205		.503			.493	
5	.464	.715			.725		
6	.310			.479			.465
7	.493	.660			.677		
8	.632	.766			.784		
9	.641	.889			.897		
10	.618	.860			.867		
11	.515	.706			.726		
12	.448	.627			.629		
13	.506		.656			.680	
14	.318	.505					.447
15	.390			.375			.316
16	.653	.942			.962		
17	.600	.870			.876		
18	.576	.823			.844		
19	.386	.508					.185
20	.259			.424			.402
21	.474	.635	.475		.648		
22	.518					.487	
23	.521	.688			.697		
24	.551	.754			.748		
25	.231		.434				.231
26	.263		.398				.316
27	.338		.531			.538	
28	.219			.504			.419
29	.169			.417			.360
30	.295			.255		.471	
31	.411		.564			.578	
32	.259		.435				.340
33	.229			.390			.338
34	.478		.725			.744	
35	.338		.552			.549	
36	.155			.393			.311
37	.277			.412			.353
38	.363		.584			.568	
39	.395		.566			.580	
40	.394		.627			.632	
41	.434		.626			.645	
42	.204			.433			.353
43	.290		.499			.503	
44	.485		.656			.688	
45	.554		.745			.757	
46	.007			.333			.266
47	.465		.639			.649	
48	.301		.421				.370
49	.257			.338		.456	
50	.121			.369			.347
No. of items	50	17	19	14	14	17	19
K-R-20	.749	.808	.705	.175	.801	.703	.161

2. Item biserials improved substantially when the internal criterion was limited to the scores on the homogeneous subset of items as compared to the case in which the criterion scores were based on the entire test.

3. The Spearman-Brown formula worked in the reverse direction. For example, the K-R-20 reliability for the entire test of fifty items was .75 for the boys. When the number of items was reduced to only a third

TABLE 2

BISERIAL CORRELATIONS AND K-R-20 RELIABILITIES FOR GIRLS' RESPONSES TO STEP-MATH 2A CLASSIFIED BY NATURE OF CRITERION SCORES

		Nature of Criterion Scores					
Item no.	Total test	Fact. I	Fact. II	Fact. III	Clus. I	Clus. II	Clus. III
1	—.007		.361				.343
2	.251			.439			.438
3	.457	.763			.762		
4	.287		.482			.480	
5	.372			.394			.395
6	.162			.320			.353
7	.389	.636			.642		
8	.479	.569			.588		
9	.466	.697			.688		
10	.478	.769			.777		
11	.469	.649			.640		
12	.280			.436			.462
13	.417		.555			.560	
14	—.010			.378			.448
15	.381	.550			.553		
16	.509	.897			.889		
17	.552	.831			.832		
18	.476	.782			.782		
19	.305	.555			.552		
20	.171			.314			.367
21	.347	.542			.539		
22	.508	.494			.600		
23	.302	.507			.495		
24	.438	.723			.729		
25	—.049			.266			.276
26	.247		.351				.346
27	.327			.377			.338
28	.280			.395			.387
29	.215			.323			.351
30	.401			.393		.446	
31	.471		.478			.579	
32	.441		.555			.553	
33	.332			.480			.444
34	.573		.629			.627	
35	.395		.517			.515	
36	.158			.309			.309
37	.460		.570			.565	
38	.363		.497			.498	
39	.483		.656			.650	
40	.417		.569			.563	
41	.383		.571			.572	
42	.211			.360			.351
43	.382		.518			.506	
44	.460		.615			.612	
45	.445		.642			.632	
46	.150			.353			.320
47	.446		.631			.621	
48	.360			.441		.446	.380
49	.283		.384			.386	
50	.209			.366			
No. of items	50	15	18	17	15	18	17
K-R-20	.709	.739	.674	.247	.745	.680	.236

of the total as in Factor I, the reliability increased to .81 though it should be dropped to .50.

4. The items which should have been rejected if analyzed against the scores on the entire test constituting the internal criterion seemed to be the ones which were essentially heterogeneous, for instance, items 1, 29, 36, 46, and 50 for boys; and items 1, 6, 14, 20, 25, 36, and 46 for girls. These items belonged to Factor III or Cluster III for which K-R-20 reliabilities had ranged from .16 to .25 only. The test resulting from the exclusion of these items still would have contained two strong, orthogonal factors or clusters and, in this sense, would have been heterogeneous. The only way in which the occurrence of more than one factor in a test can be revealed is the use of multivariate analyses.

5. Factor I and Cluster I contained almost the same items, the same applied to Factor II and Cluster II, and also Factor III and Cluster III. Thus, the choice of a particular multivariate method for analyzing matrices did not seem to be crucial.

Implications

1. Before item analysis is undertaken, it seems advisable to factor analyze the test items, especially if the test being developed is intended for wide use. However, should factor analysis not be possible in a given situation, the test developer might instead obtain the inter-item variance-covariance matrix and analyze it on a desk calculator, using the simple procedure of DuBois, Loevinger, and Gleser (1952). If the matrix contains a large number of negative and fairly large-sized covariances, the presence of more than one factor would be indicated.

2. When evidence suggests the presence of two or more strong, orthogonal factors, the test should be split into appropriate subtests, and the scores on each subtest should be used for analyzing the items included in the given subtest. Also, practical decisions should be based on subtest scores rather than on the scores for the entire test.

Reference

DuBois, P. H., Loevinger, J., and Gleser, G. C. *The construction of homogeneous keys for a biographical inventory.* Air Training Command, Human Resources Research Center, Research Bulletin 52-18; Lackland Air Force Base, San Antonio, Texas, May, 1952.

Alberta Journal of Educational Research
Vol. XIV, No. 2, June, 1968

This article presents generalizations on the adoption and diffusion of innovations. Using a general system framework, the author states these generalizations as theoretical propositions and maintains that they are applicable to systems which range in complexity from individuals to cultures. He suggests that the theory can encompass many of the theories of change and requires empirical investigation.

H. DAVID HEMPHILL
The University of Alberta

A General Theory of Innovativeness

> If there is one truth most social theorists agree on and on which they can arrive at a quick consensus, it is the lack of a viable theory of social change (Bennis, 1966, p. 99).

Despite this lack of theory, however, there has been a great deal of research on the adoption and diffusion of innovations.[1] Unfortunately, few attempts have been made to synthesize the results of this research into a set of "universally applicable generalizations."[2] The purpose of this paper is to draw from empirical evidence and place in a general system framework a *general theory of innovativeness.*

The term "general" has been chosen because of the several senses in which it applies to the theory to be developed. First, the theory is general in the sense that the concepts used are sufficiently broad to encompass a wide range of specific phenomena. Second, the theory has a foundation in empirical evidence from diverse research areas, and is applicable to systems ranging in complexity from single individuals to

---

1 For example, Ross cites over 150 studies on the adaptibility of school systems (in *Administration for adaptability.* New York: Metropolitan School Study Council, 1958); and Rogers lists more than 500 references from a number of research areas (in *Diffusion of innovations.* New York: Free Press of Glencoe, 1962).

2 Most notable among the recent attempts are those of Rogers (*Ibid.*, pp. 300-316); and D. E. Griffiths (Administrative theory and change in organizations. In Matthew B. Miles (Ed.), *Innovation in education.* New York: Bureau of Publications, Teachers' College, Columbia University, 1964. Pp. 425-436).

cultures.[3] And finally, the propositions to be stated are general in the sense that they assume that all other factors impinging on the phenomena in question are controlled either experimentally or statistically.

The General System Model

Although general system theory may not be the *sine qua non* of current behavioural theory, as some believe it to be, it does provide a useful framework and lexicon of concepts upon which to base the theory to be developed.

Von Bertalanffy (1956, p. 3) defines a *system* as a set of elements standing in interaction. Hall and Fagen (1956) expand this to define a system as a set of elements together with the relationships between the elements and their attributes. It would appear that the simpler definition is satisfactory, however, because, in the final analysis, an element is not more than a set of attributes.

All but the smallest systems have *subsystems,* and all but the largest have *suprasystems.* The environment may be defined as the set of all elements which a change of the system will affect, and in which a change will affect the system (Hall and Fagen, 1956). A system may be open or closed. *Open systems* are in continuous interaction with their environment. That is, there is a continuous flow of *input* and *output* between the system and its environment. The system-environment interaction may be in the form of exchange of material or non-material resources, such as goods, services, information or energy. Open systems tend to maintain themselves in *steady states.* A steady state is a balanced relationship among the elements of a system and its environment.

One additional concept is vital to the understanding of system theory, this is the concept of feedback. *Feedback* is the process whereby some of the output of the system is fed back into the system as a guide for the system's further functioning. This paper will use the term *negative feedback* to mean that feedback which maintains stability of the system by exerting pressure on the system to return to its original state after a slight change. *Positive feedback,* on the other hand, is pressure on the system for continued change. Miller (1965) confirms this notion when he says, "positive feedback produces continuous increments of outputs which give rise to 'spiral effects' destroying one or more equilibria of a system [p. 402]."

In order that the general system model may be applied to behavioural phenomena, a number of assumptions must be accepted. First, it must

---

[3] Boulding discusses a theoretical hierarchy of systems in terms of their complexity (in General systems theory—The skeleton of science. In Von Bertalanffy and Rapoport (Eds.), *General systems.* Yearbook of the Society for the Advancement of General Systems Theory: Vol. I, 1956. Pp. 11-17); and Miller develops hypotheses which span these levels of complexity (in Living systems: Crosslevel hypotheses. *Behavioral Science,* 1965 (October), Vol. X, 380-411).

be assumed that behavioural systems at all levels of complexity are open. Second, one must assume that these open systems maintain a steady-state, or an equilibrium, with their environment. Finally, it must be assumed that there are feedback structures within the system's environment.

Other Concepts

In addition to the concepts of the general system model, several other concepts specific to this theory must be defined.

First of all, it should be noted that *innovations* are a subset of the set of *change* phenomena. Barnett (1953) defines an innovation "as any thought, behavior, or thing that is new because it is qualitatively different from existing forms [p. 7]." Further, in this definition of innovation:

> emphasis is placed upon reorganization rather than upon quantitative variation as the criterion of a novelty (innovation). Innovation does not result from the addition or subtraction of parts. It takes place only when there is a recombination of them [p. 9].

Innovations themselves may be further divided into internal and external innovations. Internal innovations are syntheses of existing thoughts and ideas into new forms within the system whereas external innovations are those generated in the system's environment. The theory to be presented deals primarily with external innovations.

Adoption of an innovation by a system is the acceptance and implementation of the innovation within the system. Thus, *innovativeness* will be defined as the tendency of a system to adopt a given innovation or set of innovations. This tendency has been measured in terms of whether or not the system adopts, to what extent it adopts, or how early it adopts relative to other similar systems in its environment.

Finally, the concept of system *structure* must be defined. For the purposes of this discussion, structure will be defined as the demonstrated, recurrent pattern of behaviour of the system. Thus, at the individual level, the person's structure will consist of such elements as his values, habits, attitudes, beliefs, needs and goals, together with the relationships among them. The structure of a group will include an aggregate of the individual's structures together with the group's norms, roles, status hierarchy, communications networks, power and authority relations. The organization structure will be similar to that of a group with the addition of elements such as formal task and incentive structures or what Andrews and Greenfield (1966-67) call "organizational themes." At the cultural level, the structure would include such things as folkways, mores, customs, ethos, values, and ideology.

It should be noted that the structure of a system, as defined here, is a set of the system's subsystems together with their relationships. These relationships thus express dependence or independence, and subordinate

or superordinate structures of the subsystems. An example of system structure at the individual level would be a hierarchy of needs as hypothesized by Maslow (Lazarus, 1963).

Thus, the perspective to be taken is that innovativeness is related to the characteristics of the system and its interaction with the environment, and not that innovativeness is an element of the system itself, as suggested by Miles (1965).

Theoretical Propositions

With the general and specific concepts related to this theory defined, it is now possible to state a number of propositions. Following each proposition, a brief discussion of supportive research findings will be presented. It should also be noted here that the generality of the propositions necessitates that they be viewed in terms of *Ceteris Paribus*. That is, rather than compound the statements of the propositions, they should be read as if they were preceded by the phrase "all other things being equal."

Proposition 1. The greater the input of information about innovations to a system from its environment, the higher its innovativeness.

It would appear obvious that a system cannot adopt an innovation from its environment unless it has information about that innovation. Therefore, it seems reasonable to expect innovativeness to vary directly with the amount of information input.

This proposition finds support in the findings of a number of studies. At the individual level, research among farmers (Lionberger, 1960), physicians (Coleman, Katz, and Menzel, 1957), and school superintendents (Carlson, 1965) indicates that innovativeness is positively related to amount of formal education, subscription to research journals, membership in professional organizations, and visits to urban centres or professional conferences.

Further support can be found in psychological research. Using Reisman's (1961) typology of inner-versus other-directed individuals, Linton and Sofer (1959) found other-directed people were more likely to change than inner-directed ones.

At the organizational level, innovativeness has been shown as positively related to deliberate surveys of new ideas, willingness to look outside the firm, connection with universities or research institutes, and hiring of "outsiders." These relationships have been established for many types of business, industrial (Carter and Williams, 1957), and educational organizations. For example, Carlson (1962) says:

> . . . it has been demonstrated that when a school system calls for an insider, it calls for executive performance bent on system maintenance, and that when a school system calls for an outsider, it calls for executive performance bent on changing the system [p. 69].

Finally, it would not seem unreasonable to suggest that one of the main reasons for the acceleration of change in contemporary culture lies in technological advances in communication which have vastly increased the input of information to our cultural systems. Moore (1959) implies this when he suggests that "the greater the number of lines of contact, the greater the force of new influence [p. 179]."

Each of the variables which has been mentioned above may be interpreted as a measure of the quantity of input of innovation information to the system. Thus, they tend to confirm Proposition 1.

Proposition 2. The greater the perceived compatability of the innovation(s) with system structure, the higher the system innovativeness.

The first proposition dealt with the quantity of information input, whereas this one deals with the perceived quality of the innovation(s).

This proposition also enjoys the support of empirical evidence in a number of fields. For example, Fliegel (1956) found a significant relationship between innovativeness of farmers and their attitude toward new farm practices. Lindstrom (1958) found that "the adoption of practices must be preceded by the formation of favorable attitudes toward them."

Research by Hovland and others (Hovland & Janis, 1959) on persuasion and attitude change indicates that the higher the credibility of the communicator, the more likely the opinion change. The indication that ideas are accepted more readily from high prestige persons than from low, supports the proposition for, by the definition of prestige, the receiver of the idea must perceive the sender as having a structure compatible with his own, and this compatability is projected onto the innovation. Research from the same area indicates that a strong threat may be less effective than a mild threat in inducing opinion change. A mild threat is an indication of high compatibility with the individual's structure and thus positively related to innovativeness.

Stogdill (1959) expresses the notion of this proposition at the group level when he says:

> Both individuals and subgroups are likely to oppose a change which challenges the legitimacy of their roles, dislodges them from their positions, reduces the value of the outcomes they can experience in the group, imposes unnecessary hardships, or places them at a comparative disadvantage to other individuals or subgroups [p. 288].

The Junior High School study of Willower and Jones (Willower, 1962) indicated a resistance by staff members to changes which would benefit them at the expense of other members. Carter and Williams (1957) report that:

> To a great extent the receptiveness of a firm to technical change must depend on the understanding and appreciation of science and technology to be found in the various ranks of management; and this, in turn, depends on the training and experience of the individual managers [p. 66].

The classic research of Coch and French (1948) indicated that by involving staff in development of change plans, resistance to that change could be greatly reduced. These results may be interpreted as further confirmation of the proposition.

Miles (in Meierhenry, 1964) summarizes these ideas in the following statement:

> However, other things being equal, innovations which are perceived as threats to existing practice rather than mere additions to it are less likely of acceptance. More generally, innovations which can be added to an existing program without seriously disturbing other parts of it are likely to be adopted [p. 208].

The effect of cultural norms on the acceptance of innovations is illustrated by a study conducted by Pedersen (1951). He found that a Danish community accepted new agricultural practices more readily than a Polish community. Rogers (1962) suggests that the reason for this difference in readiness to accept innovations lies in the fact that:

> The Polish farmers came from a background of subsistence farming; the Danish farmers were accustomed to producing for a world food market. The cultural values of the Danish farmers facilitated the adoption of new ideas, whereas the norms of the Polish community perpetuated the *status quo* [p. 59].

One might also suggest that Weber's (1958) thesis on the relationship between the Spirit of Capitalism and the Protestant Ethic was an example of the adoption of an innovation (Capitalism) because of the compatibility of it with the structure (Protestant Ethic) of the culture.

Proposition 3. The greater the pervasiveness of the innovation within the system, the lower the system innovativeness.

Here "pervasiveness" is used in two ways. First, as the intensity of the effect on subsystems, and second, as the number of subsystems affected.

This proposition is clearly closely related to Proposition 2. However, it is possible to differentiate conceptually between a qualitative effect on system structure and a quantitative effect.

Proposition 3 suggests that an innovation which would alter an individual's habits, beliefs, and attitudes would be resisted more than an innovation which affected only his habits.

Menzel (1960) found that physicians were likely to adopt highly pervasive innovations. Lionberger (1960) summarizes additional evidence in the following way. "Generally speaking, the more complex a practice, the more change it requires in existing operations, the more slowly it will be adopted [p. 104]."

These first three propositions have dealt mainly with the input of information about innovations and their relationship to system innova-

tiveness. However, it is clear that input to one system is output from another. Therefore, innovativeness may also be conceived of as related to system output.

Proposition 4. The greater the output of the system to its environment, the lower its innovativeness.

The greater the output of a system to its environment, the stronger and more stable the system-environment interaction becomes, and thus, the greater the system's resistance to change. For, if a system has a great many "lines" of communication or interaction with its environment a change in it will upset the equilibrium or boundary-maintenance which has been established through each of these lines. If there are few of these connections, then changes will upset few system-environment equilibria.

The aforementioned study by Willower and Jones (Willower, 1962) tends to support this proposition. For instance, the resistance to changes, which would benefit some individuals at the expense of others, may be interpreted as indicating resistance related to the individual's affective output. That is, the more lines of affective output, the less innovative the individual.

Similarly, the indication by Carlson's (1965, p. 65) research that innovators tend to "know well fewer of their peers, (and to) be sought less often for advice" lends support to Proposition 4.

Gouldner (1959) discusses this notion within the concepts of interdependence and functional autonomy. He notes that "the problem of functional autonomy is of considerable significance for the analysis of tension within social systems, and thus for the analysis of social change [p. 255]." He further elaborates by saying:

> That is, those parts with least functional autonomy, those which cannot survive separation from a social system, are more likely to be implicated in its conversion than those which can.
>
> Contrariwise, those with most autonomy are most able to press for or to accept changes, when these are consistent with their own autonomy [p. 258].

Thus, sociometrically isolated individuals or isolated cultures are more susceptible to change than those which are integrated into their environments because the isolated systems have low interdependence.

The concept of "temporary systems," as developed by Miles (Miles [Ed.], 1964, pp. 454-456) has relevance here. He notes research which indicates the consequences of physical and social isolation of temporary systems. First, this isolation "tends to shear away the person's (or group's) preoccupation with, and allegiance to, "things as they are." Secondly, it reduces the role conflicts normally extant in permanent systems. Thirdly, it supplies a strong protective function. That is, "the penalties for making mistakes are reduced." Finally, the isolation of a

temporary system promotes a cohesiveness resulting from the individuals' sense of "being apart together."

If one accepts system isolation as lack of system output, then Miles' suggestion that temporary (isolated) systems change relatively easily would imply that permanent (non-isolated) systems are resistant to change. Thus, these ideas are also congruent with Proposition 4.

Proposition 5. The more the system perceives itself as compatible with its environment, the lower its innovativeness.

A system perceives itself as compatible with its environment when its needs are satisfied through inputs from the environment which are seen as resulting from its outputs to the environment. Therefore, compatibility of system and environment is a state toward which the system strives. Boulding (1956) suggests that the behaviour of a system can be explained in terms of attempts to restore these preferred states when they are disturbed by changes in the environment.

It follows, therefore, that if a system perceives itself as fairly compatible with its environment, there will be little tension or stress inducing change.[4] On the other hand, if the system and its environment are not completely compatible there will be pressure for change. Furthermore, the feedback mechanisms will bring about change directed toward increasing compatibility.

Rogers (1962) discusses research from several traditions which suggests that crisis conditions in the environment affect system innovativeness. That is, crisis conditions create changes in the environment which upset system-environment compatibility. The most recent and notable crisis to affect North American education was the launching of the Soviet sputnik. Brickell (1961) says that the rate of instructional innovation in the schools of New York State more than doubled in the 15 months following this event.

The generally inconclusive results of research on the relationship between the adoption of farm practices and farmers' values and attitudes suggests further support for the proposition. For instance, Lionberger (1960, pp. 93-94) discusses research which shows that innovativeness of farmers is related to the kinds of models with which they identify. Thus it is the system's perception of its compatibility with the environment which determines innovativeness. One would not expect particular value orientations or attitudes to predict innovativeness. Rather, the perception of the compatibility of these attitudes with those of the environment would determine whether or not the system would adopt the innovation.

---

[4] Contrary to this, McClelland *et. al.*, hypothesized that small displacements from the "adaptation level" would be favorably received by organizations (*The achievement motive*. New York: Appleton-Century-Crofts, 1953). Similarly, the work of Bexton, Heron, and Scott indicates that total compatibility of the individual with his environment may produce a state of arousal (Bexton, W. H., Heron, W., and Scott, T. H. Effects of decreased variation in the sensory environment. *Canadian Journal of Psychology*, 1954, Vol. VIII, 70-76).

Mansfield (1961) found that non-competitive firms were slower to innovate than competitive ones. Carlson (1965, pp. 6-7) calls these non-competitive organizations "domesticated." If one interprets lack of competition as compatibility with the environment, then this notion confirms the proposition.

March and Simon (1959) explain these phenomena in the following way:

> Individuals and organizations give preferred treatment to alternatives that represent continuation of programs over those that represent change . . . persistence comes about primarily because the individual or organization does not search for or consider alternatives to the present course of action unless that present course is in some sense "unsatisfactory" [p. 173].

Proposition 6. The greater the pervasiveness of the output of a system, the lower its innovativeness.

Again, "pervasiveness" is used as both the intensity of effect on suprasystems, and as the number of suprasystems affected. This meaning of pervasiveness is very similar to that used by Etzioni (1961). He says, "the range of pervasiveness is determined by the number of activities in or outside the organization for which the organization sets norms [p. 163]."

Rogers (1962, p. 177) notes several research projects which indicate that more specialized sytems tend to be more innovative. It would seem reasonable to suggest that specialization of a system could be considered an indicator of low pervasiveness of output, and thus this research tends to support Proposition 6.

The propositions which have thus far been presented have considered input and output of the system separately. It is obvious, however, that for any open system, the two are interdependent. The concept of feedback connects system input to output, and suggests a final proposition.

Proposition 7. The greater the ratio of positive to negative feedback of the system, the higher the system innovativeness.

Although systems tend toward steady states, it is clear that systems do change. The changes must, therefore, be a result of variation of input. If a system is in *stationary equilibrium* the negative feedback is sufficiently great to overcome any positive feedback, and the system will return to its original state despite a change in input. If, however, the positive feedback is greater than the negative feedback, the system will not return to its original state after a change of input, but will achieve a different equilibrium. A system continuously undergoing this type of process is said to be in *dynamic equilibrium* (Chin, 1961). In the terminology used here, a system in dynamic equilibrium would be more innovative than a system in stationary equilibrium.

Although positive and negative feedback have not been researched *per se,* the wealth of studies at Columbia University has shown that the adaptability (a synonym for innovativeness) of school systems is very largely a function of the characteristics of the population of the community. Mort (1964) summarizes this research as follows:

> Communities vary in the degree to which they take on new practices. Indications are that this is a community characteristic. A community that is slow to adopt one innovation tends to be slow to adopt others. A pioneer in one area tends to be a pioneer in other areas.
>
> Explanation of the difference in educational adaptability of communities can be found in no small degree in the character of the population, particularly in the level of the public's understanding of what schools can do, and citizens feeling of need for education for their children. This appears to set the posture of the community toward financial support, and toward what teachers are permitted to do—and tends to shape the staff by influencing personnel selected and kept in the community [p. 326].

To the extent that community understanding, financial support and permissiveness can be considered as representative of the ratio of positive to negative feedback, this research tends to support the proposition.

Predominance of negative feedback would, according to the proposition, result in low innovativeness. This phenomenon was noted in a study reported by Eichholz and Rogers (1964). They found that ninety-three per cent of the teachers in their sample had rejected or discontinued the use of electromechanical innovations through experience. That is, the experience of the teacher with the innovation had resulted in negative feedback sufficient to mitigate against its future use.

Characteristics of Innovation

The theory of innovativeness presented here would appear to suffer from one serious omission. At no time has there been direct consideration given to the characteristics of the innovation. This omission was not unintentional. Characteristics of innovations which have been shown to be related to innovativeness appear to be closely associated with the propositions presented.

Rogers (1962, pp. 124-134) suggests that the (*a*) relative advantage, (*b*) compatibility, (*c*) complexity, (*d*) divisibility, and (*e*) communicability of an innovation affect its adoption. These characteristics would seem to be implied by, if not included in, the stated propositions. The communicability of the innovation is clearly implied in Proposition 1. The complexity and divisibility of the innovation are obviously related to the extent of change in the system required for adoption, and therefore, implied by Proposition 3. Finally, the compatibility and relative advantage of the innovation represent the compatibility of the innovation with the structure of the system, Proposition 2.

Conclusions

It is clear that the propositions stated here have been discussed as relatively discrete relationships. However, they are obviously statements of interdependent concepts and relationships. Of particular interest is the paradoxical symmetry of Propositions 1 to 3, and 4 to 6. The second triad hypothesizes relationships between certain output concepts and innovativeness which are to a certain extent converse to those implied in the first triad. For instance, Proposition 1 suggests that innovativeness rises when innovation information input rises, whereas Proposition 4 states that the greater the output from a system to its environment the lower its innovativeness. And yet the assumption of feedback structures means that there is a necessary and direct connection between system output and system input. If this is the case, then the relationships proposed in the first triad will tend to negate the relationships in the second triad, and the system will remain static. The result is basic system resistance to innovation.

Nevertheless, systems do change, and they do adopt innovations. Therefore, the relationships hypothesized must have differential magnitudes which allow innovativeness. It is suggested that the theory presented here is capable of accounting for the "orderly reciprocity of persistence and change (Williams, 1966, p. 15)."

Further, it is suggested that this theory is sufficiently general in nature to encompass the notions and theories of change in the disciplines of psychology, social psychology, sociology, and anthropology. For instance, at the individual level. Festinger's (1957) theory of cognitive dissonance is closely related to Proposition 1. At the group level, the principles of change suggested by Cartwright (1951) may be interpreted in terms of these propositions.

Theories of acculturation or cultural innovativeness may also be accounted for by the present theory. For instance, Toynbee's (1933) concept of "challenge and response" in societal development is clearly related to Proposition 5. Barnett (1953) states these hypotheses:

> . . . an individual will not accept a novelty unless in his opinion it satisfies a want better than some existing means at his disposal.
>
> * * *
>
> . . . people develop tastes and preferences under the influence of particular experiences, and these orientations are significant for the acceptance or rejection of new ideas [pp. 378-379].

These appear to be specific cases of Proposition 2.

Similarly, Moore (1959) notes a number of acculturative situations of readiness and capability for novelty (innovativeness in the terms used here).

1. The greater the number of lines of contact with an external culture, the greater is the force of the new influence.

2. Acceptance of novelty is much greater and easier if contact with the new element is continuous and, in principle, permanent.

3. Under equivalent or analogous conditions simple elements are more readily transferable than complex ones and simple cultural forms are more transferable than complex cultural meanings.

4. Acceptance is greater and easier if congruent with the sense of advantage of those with recognized vested interests.

5. The higher the prestige level of the innovators or donors, the faster the rate of acceptance.

6. The rate and areas of acceptance of novelty are a function of the intensity and the directions of prevailing expectations of change.

7. Acceptance varies with the tension or gratification level which characterizes relationships generally between the innovator-donor and the accepter cultures [pp. 179-198].

It is clear that one could also place these seven statements in the framework of propositions on the general theory developed here, as one could many of the theories of learning, role behaviour, leadership, and power and authority.

In addition to its generality in terms of previous theories, this theory appears to present a framework upon which many of the research findings on the adoption and diffusion of innovations may be placed. There is no doubt, however, that the concepts and relationships which have been presented here require confirmation through further empirical investigation. It is the writer's contention that these concepts and propositions can be "operationalized" and studied in diverse situations. For, as Getzels (1958) suggests, "in the long run, the best test of the usefulness of a theoretical formulation is whether or not it generates research and inquiry [p. 150].

References

Andrews, J. H. M., and Greenfield, T. B. Organizational themes relevant to change in schools. *Ontario Journal of Educational Research,* Vol. IX, No. 2, 81-99.

Barnett, H. G. *Innovation: The basis of cultural change.* New York: McGraw-Hill, 1965.

Bennis, W. G. *Changing organizations.* New York: McGraw-Hill, 1966.

Boulding, K. E. General systems theory—the skeleton of science. In L. Von Bertalanffy, and A. Rapoport (Eds.), *General Systems.* Yearbook for the Society for the Advancement of General Systems Theory, 1956, Vol. I, 18.

Brickell, H. M. *Commissioner's 1961 catalog of educational change.* Albany, New York: State Education Department, 1961. P. 27.

Carlson, R. O. *Executive succession and organizational change.* Chicago: Midwest Administration Center, The University of Chicago, 1962. P. 69.

Carlson, R. O. *Adoption of educational innovations.* Eugene, Oregon: Center for the Advanced Study of Educational Administration, The University of Oregon, 1965.

Carter, C. F., and Williams, B. R. *Industry and technical progress.* London: Oxford University Press, 1957. Pp. 178-180.

Cartwright, D. Achieving change in people: Applications of group dynamics theory. *Human Relations,* 1951, Vol. IV, 388-391.

Chin, R. The utility of system models and developmental models for practitioners. In W. G. Bennis, K. D. Benne, and R. Chin (Eds.), *The planning of change.* New York: Holt, Rinehart & Winston, 1961. Pp. 201-214.

Coch, L., and French, J. R. F., Jr. Overcoming resistance to change. *Human Relations,* 1948, Vol. I, 512-532.

Coleman, J., Katz, E., and Menzel, H. The diffusion of an innovation among physicians, *Sociometry,* 1957, Vol. XX, 253-270.

Eicholz, G., and Rogers, E. M. Resistance to the adoption of audio-visual aids by elementary school teachers: Contrasts and similarities to agricultural innovations. In M. B. Miles (Ed.), *Innovation in education.* New York: Bureau of Publications, Teachers' College, Columbia University, 1964. Pp. 299-316.

Etzioni, A. *A comparative analysis of complex organizations.* New York: The Free Press of Glencoe, 1961. P. 163.

Festinger, L. *A theory of cognitive dissonance.* Evanston, Illinois: Row, Peterson, 1957.

Fliegel, F. C. A multiple correlation analysis of factors associated with adoption of farm practices. *Rural Sociology,* 1956, Vol. XXI, 284-292.

Getzels, J. W. Administration as a social process. In A. W. Halpin (Ed.), *Administrative theory in education.* Chicago: The Midwest Administration Center, The University of Chicago, 1958. P. 150.

Gouldner, A. W. Reciprocity and autonomy in functional theory. In L. Gross (Ed.), *Symposium on sociological theory.* New York: Harper & Row, 1959. P. 255.

Hall, A. D., and Fagen, R. E. Definition of system. In L. Von Bertalanffy, and A. Rapoport (Eds.), *General Systems.* Yearbook for the Society for the Advancement of General Systems Theory, 1956, Vol I, 18.

Hovland, C. I., and Janis, F. I. (Eds.), *Personality and persuability.* New Haven, Connecticut: Yale University Press, 1959.

Lazarus, R. S. *Personality and adjustment.* Englewood Cliffs, N.J.: Prentice-Hall, 1963. P. 67.

Lindstrom, D. E. Diffusion of agricultural and home economics practices in a Japanese rural community. *Rural Sociology,* 1958, Vol. XXIII, 174.

Linton, H., and (Sofer) Graham, E. Personality correlates of persuability. In C. I. Hovland, and F. I. Janis (Eds.), *Personality and persuability.* New Haven, Connecticut: Yale University Press, 1959. P. 70.

Lionberger, H. *Adoption of new ideas and practices.* Ames, Iowa: The Iowa State University Press, 1960.

Mansfield, E. Technical change and the rate of imitation. *Econometrica,* 1961, Vol. XXIX, 741-766.

March, J. G., and Simon, H. A. *Organizations.* New York: Wiley, 1959. P. 173.

Menzel, H. Innovation, integration and marginality: A survey of physicians. *American Sociological Review,* 1960, Vol. XXV, 704-713.

Miles, M. B. Educational innovations: Some generalizations. In W. C. Meierhenry (Ed.), *Media and educational innovation.* Lincoln, Nebraska: University of Nebraska Press, 1964. P. 208.

Miles, M. B. (Ed.), *Innovation in education.* New York: Bureau of Publications, Teachers' College, Columbia University, 1964. Pp. 425-436.

Miles, M. B. Planned change and organizational health: Figure and ground. In R. O. Carlson *et. al., Change processes in the public schools.* Eugene, Oregon: Center for the Advanced Study of Educational Administration, University of Oregon, 1965. Pp. 11-34.

Miller, J. G. Living systems: Cross-level hypotheses. *Behavioral Science,* 1965, (October), Vol. X, 380-411.

Moore, W. F. *Industrialization and labor: Social aspects of economic development.* Ithica, New York: Cornell University Press, 1959. P. 179.

Mort, P. R. Studies in educational innovation from the institute of administrative research: An overview. In M. B. Miles (Ed.), *Innovation in education.* New York: Bureau of Publications, Teachers' College, Columbia University, 1964. P. 326.

Pedersen, H. A. Cultural differences in the acceptance of recommended practices. *Rural Sociology,* 1951, Vol. XVI, 37-49.

Riesman, D., *et. al. The lonely crowd.* New Haven, Connecticut: Yale University Press, 1961.

Rogers, E. M. *Diffusion of innovations.* New York: The Free Press of Glencoe, 1962.

Stogdill, R. M. *Individual behavior and group achievement.* New York: Oxford University Press, 1959. P. 288.

Toynbee, A. J. *The study of history.* London: Oxford University Press, 1933. 12 Vols.

Von Bertalanffy, L., and Rapoport, A. (Eds.), *General Systems.* Yearbook for the Society for the Advancement of General Systems Theory, 1956, Vol. I.

Weber, M. The protestant ethic and the spirit of capitalism, (trans. Talcott Parsons). New York: Scribner, 1958.

Williams, T. R. The study of change as a concept in cultural anthropology. *Theory into Practice,* 1966, Vol. V, 15.

Willower, D. J. Barriers to change in educational organizations. *Theory into Practice,* 1962, (December), Vol. II, 258.

Alberta Journal of Educational Research
Vol. XIV, No. 2, June, 1968

This article summarizes evidence of the construct validity of several "culture-reduced" measures of intellectual ability for natives. Three characteristics of such tests suggest further research. An additional conclusion is that large proportions of Canadian native pupils of early school age have the general intellectual ability which seems necessary to participate fully in the larger Canadian community.

R. S. MacARTHUR
The University of Alberta

Assessing Intellectual Potential of Native Canadian Pupils: A Summary

Current emphasis on the development of human resources has underlined the need for valid estimates of the intellectual potential of individuals from a variety of backgrounds to aid in adapting teaching or training to that potential. Now that some long-term predictive validity data are available, this article summarizes evidence from a series of studies of the construct validity of several "culture-reduced" measures of general intellectual ability for native pupils of the Canadian West and North.

Rationale

A suggested rationale to guide such studies has been outlined elsewhere by MacArthur (1968) and by West and MacArthur (1964). In brief, it conceptualizes abilities as organized in a hierarchy from relatively specific abilities at the bottom to general intellectual ability at the top, resembling the models of Vernon (1965) and of Gagné (1962). The development of these abilities takes place through a sort of cumulative transfer as innate predispostions interact with environmental conditions, along lines suggested by Hebb (1949) and by Piaget (1964). Since environmental conditions may differ considerably from one culture to another, so may the patterning and nature of abilities at all levels of the hierarchy, as suggested by Ferguson (1954) and by Biesheuval (1962).

Abilities high in the hierarchy are less affected by particular environmental experiences, and hence measures of general intellectual ability with minimum cultural bias should provide least bad estimates of intellectual potential for individuals from other cultures.

This line of reasoning has led to eight criteria suggested for the construct validity of a measure of general intellectual potential for pupils who are not from middle-class urban European-American backgrounds.

1. It should largely sample the broad factor of general intellectual ability running through a variety of European-American kinds of intellectual tasks, including reading, linguistic, and arithmetic achievement.

2. It should show less difference between cultures in cross-cultural administration than do alternative measures of intelligence. This is the concept of a "culture-reduced" test (as contrasted with the impractical concepts of "culture-free" or "culture-fair"), in the sense of one that minimizes but does not eradicate the effects of bias in terms of a particular cultural environment.

3. It should show moderate relationship with concurrent school achievement or trade efficiency.

4. It should show long-term validity as a predictor of success in intellectual tasks when appropriate adaptive intervening treatments have been employed.

5. It should be reliable.

6. It should show stability under changed environmental conditions, relative to that of alternative measures of intelligence.

7. It should minimize effects of test sophistication, providing plenty of appropriate practice experience, having directions depending little if at all upon language, and be unspeeded.

8. It should be administratively practical in that it has face validity and intrinsic interest, it is economical, it can be administered to groups by competent teachers, and it has appropriate norms or is amenable to the production of such norms.

Data Collection

Samples

Data are here summarized for three groups of samples. Group I consisted of 126 Metis pupils in Grades 1 to 8 at Faust, Alberta, with replication on 155 Metis and Indian pupils in Grades 1 to 9 at Fort Simpson, N.W.T. These pupils were sub-divided into eight samples for analysis by grade-levels, and are described in MacArthur (1962). Group II, the prediction samples, consisted of the 45 Metis pupils who were in Grades 1 to 3 in the above Faust samples and were attending school at Faust four years later, and of 97 Eskimo pupils in two groups ages 6 to 9 and 9½ to 12 years, at Inuvik and Tuktoyaktuk, for whom three-year predic-

tion data were available. These samples are described in Rattan and MacArthur (1968). Group III, the norming samples, consisted of 792 Indian-Metis, 510 Eskimo, and 709 Whites who were representative of pupils attending Grades 1 to 8 of the schools of the Mackenzie District, N.W.T. The samples are described in MacArthur (1965).

Measures

Among the group ability tests administered to the samples of Group I were those listed in Table 1, the first five of which were hypothesized to be somewhat culture-reduced measures of general intellectual ability, and the others of which include conventional measures of scholastic aptitude and school achievement. The actual tests used, with levels considered appropriate for the respective grade-levels, were: Progressive Matrices (Coloured and Standard), Safran Culture-Reduced Intelligence Test Scale 1, IPAT Cattell test of g Scale 2, Lorge-Thorndike Non-Verbal Intelligence Test (four levels with three sub-tests and a total score for each level), California Short-Form Test of Mental Maturity (three levels with six sub-tests and a Total score for each level), Detroit Beginning First-Grade Intelligence Test, Otis Quick-Scoring Mental Ability Tests (Alpha and Beta), and California Achievement Battery (four levels with three sub-tests and a Total score for each level). A selection from these tests was administered to the prediction samples of Group II and the norming samples of Group III.

Results and Discussion

For each sample of Group I separately, after normalizing most scores, the intercorrelations of the variables listed in Table 1 were principal-factor analyzed, with unities inserted in the diagonal cells and common factors whose corresponding eigenvalues were greater than one considered significant. Where no entry for a test occurs, that test was not administered to the particular sample. The unrotated first principal factor loadings shown in Table 1 may be considered to indicate the degree to which a test is measuring, for a sample, the most important single intellectual ability factor running through these cognitive tasks which include conventional verbal intelligence tests and school achievement tests. The communalities shown in Table 2 may be considered to indicate the degree to which a test is measuring, for a sample, all of the various common factors running through the battery of ability and school achievement tests. In later studies similar results were obtained for several Eskimo samples. Tables 1 and 2 taken together indicate the degree to which the various measures are meeting Criterion 1 for culture-reduced measures of general intellectual ability. (Rotated factor patterns are reported in MacArthur (1962). Other factor analyses were made using sub-tests of the Lorge-Thorndike and California Mental Maturity Tests.)

TABLE 1

UNROTATED FIRST PRINCIPAL FACTORS
Metis and Indian Samples

Sample	Gr. 7 & 8		Gr. 5 & 6		Gr. 2 & 3		Gr. 1	
	Faust	Simpson	Faust	Simpson	Faust	Simpson	Faust	Simpson
N	23	32	29	58	42	46	32	19
Mean Age	14-4	13-10	12-10	13-11	9-4	10-2	7-9	7-6
Prog. Matrices	.76	.59	.88	.51	.82	.46	.64	.76
SCRIT	.55	.55	.81	.56	.79	.67	.73	.71
IPAT Cattell	.75	.55	.78	.69				
Lorge-Th. NV Tot.	.77	.86	.89	.82	.76	.73	.75	.73
CTMM Non-Lang.	.60	.32	.34	.57	.75	.67	.36	.88
CTMM Lang.	.86	.79	.76	.71	.83	.62	.58	.92
Detroit Begin.					.63		.49	.84
Otis Beta	.79	.80	.63	.88				
Otis Alpha NV						.71		
Otis Alpha Vbl.						.73		
TOGA Total	.70		.82		.47		.68	
Calif. Ach. Rdg.	.82	.87	.60	.86	.90	.74	.69	.94
Calif. Ach. Arith.	.80	.83	.62	.90	.80	.72	.81	.91
Calif. Ach. Lang.	.82	.87	.07	.81	.86	.78	.75	.68
Prop. Total Variance	.57	.53	.49	.55	.59	.48	.43	.67

TABLE 2

COMMUNALITIES
Metis and Indian Samples

Sample	Gr. 7 & 8		Gr. 5 & 6		Gr. 2 & 3		Gr. 1	
	Faust	Simpson	Faust	Simpson	Faust	Simpson	Faust	Simpson
No. Common Factors	3	3	4	3	3	3	3	2
Prog. Matrices	.72	.64	.80	.66	.89	.75	.51	.71
SCRIT	.91	.69	.79	.67	.80	.67	.57	.90
IPAT Cattell	.71	.77	.78	.66				
Lorge-Th. NV Tot.	.74	.77	.83	.82	.60	.65	.57	.55
CTMM Non-Lang.	.83	.94	.91	.87	.67	.71	.94	.76
CTMM Lang.	.86	.87	.68	.73	.71	.74	.63	.88
Detroit Begin.					.77		.81	.77
Otis Beta	.87	.87	.82	.84				
Otis Alpha NV						.59		
Otis Alpha Vbl.						.67		
TOGA Total	.73		.75		.87		.47	
Calif. Ach. Rdg.	.80	.89	.95	.87	.90	.75	.77	.88
Calif. Ach. Arith.	.79	.70	.75	.84	.87	.61	.73	.92
Calif. Ach. Lang.	.84	.77	.91	.75	.87	.85	.77	.63
Prop. Total Variance	.80	.79	.82	.77	.79	.70	.68	.78

For the tests for which entries appear in Table 3, scores were obtained on samples totalling 780 pupils representative of the Grade 7 pupils in Edmonton, and the Grades 6, 3, and 1 pupils of Calgary; these were converted to T-scores based on the respective White samples, so that the means for each White urban sample on each test all were 50 with standard deviation 10. Scores for the eight Indian-Metis samples of Group I were then expressed in terms of the most nearly appropriate White urban T-scores, with means for the natives as shown in Table 3. This table, the data of which are extracted from West and MacArthur (1962), thus shows the degree to which the various measures are meeting Criterion 2 (the tests with highest means in any column have the least cultural bias). It provides an empirical definition of the concept "culture-reduced", for these Metis and Indian samples.

TABLE 3

MEAN T-SCORES FOR METIS AND INDIAN SAMPLES
BASED ON EDMONTON OR CALGARY SAMPLES
(White urban means are 50, S.D. of 10, in each instance)

Sample	Gr. 7 & 8 in Edmonton T		Gr. 5 & 6 in Calgary T		Gr. 2 & 3 in Calgary T		Gr. 1 in Calgary T	
	Faust	Simpson	Faust	Simpson	Faust	Simpson	Faust	Simpson
1. Prog. Matrices	41	45*	37*	40*	35*	36*	41*	41*
2. SCRIT			36*	39*	38*	40*	49*	41*
3. IPAT Cattell	41	41*	38*	39*				
4. Lorge-Th. NV Tot.	42*	45*	31	27	33	24	39*	35*
5. Lorge-Th. NV Pt. 3	43*	46*	34	29	45*	37*	43*	40*
6. Lorge-Th. NV Pt. 2	41	45*	34	29	41*	40*	45*	39*
7. Lorge-Th. NV Pt. 1	47*	45*	37*	34*	29	20	38	32
8. CTMM Non-Lang.	43*	41*			41*	42*		
9. CTMM Spatial	46*	45*			48*	53*		
10. CTMM Logical	38	35			35*	30*		
11. CTMM Numerical	37	36			32	26*		
12. CTMM Verbal	36	33			28	17		
13. CTMM Lang.	33	30			26	17		
14. CTMM Total	36	31			29	21		
15. Detroit Begin.							35	27
	*Sig. Diff. from CTMM Total at .01 level		*Sig. Diff. from L-Th NV Total at .01 level		*Sig. Diff. from CTMM Total at .01 level		*Sig. Diff. from Detroit Begin. at .01 level	

Using median values over the eight samples of Group I, Table 4, which represents the core of this article, presents a summary evaluation of nine tests which were hypothesized to be culture-reduced measures of general intellectual ability in terms of: Criterion 1—(a) unrotated first principal factor loadings, and (b) communalities; Criterion 2 — mean T-scores based on White urban standards; Criterion 3 — correlation with concurrent school achievement as measured by California Achievement Tests Total. The crudity of such overall summaries using median values,

not always over the same numbers of samples, is emphasized; for detail the reader is referred to previous tables and to references already given. In general, however, it is seen that tests least badly meeting the first three criteria set up for measures of intellectual potential with minimum

TABLE 4

SUMMARY EVALUATION OF NINE CULTURE-REDUCED TESTS IN TERMS OF FOUR CRITERIA

Criterion	1		2	3	4		
					Long-term r with Schl. Ach.		
	Mdn. 1st Unrotated Factor	Mdn. Communality	Mdn. of Mean T-scores	Mdn. r Concurrent Schl. Ach.	45 Metis over 4 yrs.	62 Esk. over 3 yrs.	37 Esk. over 3 yrs.
1. Prog. Matrices*	.70	.72	41	.53	.57	.43	.60
2. SCRIT*	.69	.74	40	.49	.62		
3. IPAT Cattell*	.72	.74	40	.43			
4. Lorge-Th. NV Tot.	.77	.70	34	.64			
5. Lorge-Th. NV Pt. 3	.65	.81	42	.51			
6. Lorge-Th. NV Pt. 2	.71	.74	41	.57			
7. Lorge-Th. NV Pt. 1	.60	.72	36	.48			
8. CTMM Non-Lang.	.59	.85	42	.41	.50		
9. CTMM Spatial	.43	.78	47	.29	.37		
10. CTMM Logical	.63	.71	35	.59	.59		
11. CTMM Numerical	.66	.80	34	.58	.52		
12. CTMM Verbal	.72	.82	31	.73	.54		
13. CTMM Lang.	.78	.73	28	.73	.68		
14. CTMM Total			30	.72	.64		
15. Detroit Begin.	.63	.79	31	.44			
16. Otis Beta	.79	.85		.76			

*Tests least badly meeting these criteria for construct validity.

cultural bias are Progressive Matrices, Safran Culture-Reduced Intelligence Test, Cattell test of g Scale 2, and some sub-tests of the Lorge-Thorndike Non-Verbal Intelligence Tests. The Non-Language and Spatial sub-tests of the California Test of Mental Maturity, though showing less cultural bias than other measures in the battery, are not as highly loaded on the first principal factor and do not relate as highly to school achievement. As might be expected, the conventional tests 10 to 16 of Table 4 are somewhat more highly related to school achievement, but show much more cultural bias against the Metis and Indian pupils.

Table 4 also summarizes, for some of the measures, data from Rattan and MacArthur (1968) with respect to Criterion 4 — long-term predictive validity for samples of Group II. For the 45 Metis in the prediction sample, Coloured Progressive Matrices and SCRIT did not differ significantly from any of the more conventional tests 10 to 14 in predicting California Achievement Total over a four-year period.

Reliabilities of the various tests were not studied directly for these particular samples. However, communalities provide minimum estimates

of reliabilities (reliabilities must be as high as, and may be much higher than, communalities). In general, all of the measures studied would appear to have satisfactory reliabilities. Stability coefficients for Progressive Matrices and SCRIT, for the Metis prediction sample, did not differ significantly over the four-year period from those of the conventional CTMM Total or Verbal, Numerical and Logical sub-tests. Directions for the two former tests rely heavily on gesture and demonstration, and their time limits were so set that all but the very slow pupils were able to finish; however, more experimental work with respect to Criterion 7 would be desirable.

All of the culture-reduced tests studied appear to be intrinsically interesting to the Metis, Indian, and Eskimo pupils, are relatively economical, and can be readily administered as group tests by competent teachers. Appropriate norms for such pupils often present a problem; on the basis of the above and similar evidence, the Canadian Department of Northern Affairs undertook the Mackenzie District Norming Project, the report of which by MacArthur (1965) presents age norms for Eskimo, Indian-Metis, and Whites separately, based on samples of Grades 1 to 8 pupils representative of the whole of the Mackenzie District of the North West Territories, for Coloured and Standard Progressive Matrices, Otis Alpha Non-Verbal, and Lorge-Thorndike Non-Verbal Levels 3 and 4.

Conclusions

1. The evidence summarized above would seem to suggest that, until much-needed research can improve on them, such culture-reduced measures of general intellectual ability as Progressive Matrices and SCRIT for all grades 1 to 8, and IPAT Cattell and Lorge-Thorndike Non-Verbal for the higher of these grades, should be included in testing programmes aimed at helping us understand and adapt instruction to the abilities and potential of individual native Canadian pupils. However, conditions and pupils are changing rapidly these days, and construct validities may change similarly. Continuing studies such as those summarized are needed.

2. One might ask—what is there about such tests as Progressive Matrices which helps them, at least for our samples, meet these criteria taken together less badly than do conventional group so-called intelligence tests? Three replies suggest themselves: (a) the items form something of an age-scale sampling stages in the development of human cognition, starting with perception-dominated items, and proceeding through reversible concrete operations, to propositional or formal operations; (b) they use as stimuli symbols, which, though dependent on learning, are likely to be learned in a variety of cultures; (c) arrangement of items in the test itself forms a crudely-programmed sample of learning-on-the-

spot. These suggest three directions in which continued basic research should help us with the practical problems of assessing the general intellectual potential of individual pupils or candidates from other cultures.

3. If the validity data presented here for Progressive Matrices are accepted, an additional conclusion from the Mackenzie District Norming Project is that large proportions of Canadian native pupils of early school age *have* the general intellectual ability to participate fully in the larger Canadian community. On this test, for seven-year-olds, for example, the 90th percentiles for Eskimo, Indian-Metis, and Whites were respectively 25, 25, 29; the 50th percentiles were 18, 16, 20.5; the 10th percentiles were 13, 9, 14. For nine-year-olds, the 90th percentiles for Eskimo, Indian-Metis, and Whites were respectively 31, 31, 33; the 50th percentiles were 22, 21, 26; the 10th percentiles were 14, 15, 18. Thus, of these young native pupils, many scored well above the average Whites. At ages 11 and 13 years few native pupils scored above average Whites, but many scored well above the 10th percentile for Whites.

These culture-reduced tests are still considered to be culturally biased when used in this context. Nevertheless these and other data from the references herein cited clearly indicate very considerable intellectual potential among the early-school-age native pupils of the Canadian West and North.

References

Biesheuval, S. The detection and fostering of ability among underdeveloped peoples. *Yearbook of Education,* 1962, 337-353.

Ferguson, G. A. On learning and human ability. *Canadian Journal of Psychology,* 1954, Vol. VIII, 95-112.

Gagné, R. M. The acquisition of knowledge. *Psychological Review,* 1962, Vol. XXIX, 355-365.

Hebb, D. O. *The organization of behavior.* New York: Wiley, 1949.

Inhelder, B., and Piaget, J. *The early growth of logic in a child.* New York: Harper & Row, 1964.

MacArthur, R. S. Assessing the intellectual ability of Indian and Metis pupils at Fort Simpson. Ottawa: Department of Northern Affairs and National Resources, 1962.

MacArthur, R. S. Mackenzie norming project. Ottawa: Department of Northern Affairs and National Resources, 1965.

MacArthur, R. S. Some differential abilities of Northern Canadian native youth. *International Journal of Psychology,* 1968, Vol. III.

Rattan, M. S., and MacArthur, R. S. Longitudinal prediction of school achievement for Metis and Eskimo pupils. *Alberta Journal of Educational Research,* 1968, Vol. XIV, No. 1.

Vernon, P. E. Ability factors and environmental influences. *American Psychologist,* 1965, Vol. XX, 723-733.

West, L. W., and MacArthur, R. S. An evaluation of selected intelligence tests for two samples of Metis and Indian children. *Alberta Journal of Educational Research,* 1964, Vol. X, No. 1, 17-27.

Alberta Journal of Educational Research
Vol. XIV, No. 2, June, 1968

Ziller's disguised measure of test-risk is compared with a reported measure of test-risk. Correlations are reported for the two measures, and the distributions are compared. The evidence presented supports the disguised measure as a valid assessment of risk taking.

MALCOLM J. SLAKTER and AUBREY H. RODEN
State University of New York at Buffalo

Comparison of a Disguised and Reported Measure of Risk Taking On Objective Examinations

In an objective achievement or aptitude examination, the examinee is confronted with a series of items, each of which contains a stem and a set of alternative responses. For each item, the examinee is directed to select the single best alternative for the given stem. In an attempt to diminish the effects of guessing upon test score, the examinee is oftentimes further instructed to omit those items to which he does not know the answer. The latter instruction comes in many forms and varieties, but all belong to the set of do-not-guess directions. Do-not-guess directions are generally accompanied by a "correction formula" or penalty for incorrect responses; i.e., test score equals the number of correct responses minus some fraction of the number of incorrect responses. Therefore, under do-not-guess directions, it is clear that whenever the examinee encounters an item in which he cannot select the "best" alternative with certainty, there exists a risk taking situation. If the decision is to omit the item, the use of the "correction formula" will not affect the test score. On the other hand, if the decision is to select one of the alternatives, the use of the "correction formula" will increase or decrease the test score, depending on whether or not the alternative selected is the best alternative.

There is ample evidence that when confronted with these risk taking situations, examinees differ in their risk taking propensities (Slakter,

1967a, 1967b; Stone, 1962; Swineford, 1938); i.e., on a given examination, examinees differ reliably in their tendency to omit items or to guess. It is reasonable to assume that this particular behaviour is an expression of a more general motivational pattern. Indeed, there is evidence that risk taking on objective examinations is related to behaviours such as vocational choice (Ziller, 1957a), dominance-submission (Votaw, 1936), and maladjustment (Sherriffs and Boomer, 1954).

In addition, there is evidence that the low risk taker on examinations tends to be penalized by his reluctance to respond to items that are not known with certainty (Sherriffs and Boomer, 1954; Slakter, 1967a; Votaw, 1936). These studies have demonstrated that when low risk takers are forced to respond to all items, their test score will generally increase even though the usual penalty for guessing is applied. Therefore, these findings provide evidence that the personality trait of risk taking confounds the achievement or aptitude being measured by the examination.

There have been several disguised measures of test-risk proposed in the literature (Danielson, 1956; Slakter, 1967a, 1967b; Swineford, 1938), but the most widely used to date has been that developed by Ziller (1957a). Ziller's index is a function of the number of incorrect and omitted items, and the number of alternatives per item. Specifically, each examinee receives a risk taking score R, where

$$R = \frac{\frac{n}{n-1} w}{\frac{n}{n-1} w + u},$$

and n is the number of alternatives for each item, w is the number of incorrect responses for the examinee, and u is the number of items unattempted by the examinee. For each subject, R is essentially an estimate of the ratio of items guessed to items not known. The range of R is from 0 to 1, with low values of R indicating low risk, and high values indicating high risk.

Another method of determining the extent of risk taking by an examinee is simply to ask him; i.e., a self-report. Naturally, this self-report suffers from the deficiencies of all self-reports: (*a*) the examinee must know the answer to the question posed, and (*b*) assuming he knows the answer, he must be willing to share the information with the person administering the self-report. Unfortunately, there is evidence that examinees are unable to clearly distinguish between those items on which they have complete, partial, or no information (Granich, 1931; West, 1923); i.e., examinees do not know precisely the extent of their test-risk. However, the lack of precise knowledge of degree of test-risk does

not necessarily imply that there is no knowledge. Clearly, the construct validity of the Ziller measure would be enhanced if there existed strong positive relations between the disguised Ziller test-risk score and the reported test-risk score. The main problem of interest in this paper, therefore, was to compare the Ziller index with a reported measure of test-risk. Specific areas of interest were (*a*) the correlation between the disguised and reported measures, and (*b*) a comparison of the distributions of the disguised and reported measures.

Method

The examination used in this study was the Concept Mastery Test (Terman, 1950) or CMT. The do-not-guess directions of the CMT instructs examinees to: "Omit those items that you could answer by pure guess, but answer all you *think* you know, even if you are not quite certain." The items in the first part of the CMT have two atlernatives each, and are based upon vocabulary; the items in the second part of the CMT have three alternatives, and are based upon verbal reasoning. Since the two parts of the CMT have different numbers of alternatives, a slight modification is necessary in the calculation of R, the disguised measure. Therefore, in this study

$$R = \frac{2\,w_1 + \frac{3}{2}\,w_2}{2\,w_1 + \frac{3}{2}\,w_2 + u},$$

where w_1 is the number of incorrect responses on part 1, and w_2 is the number of incorrect responses on part 2.

Immediately following the administration and collection of the CMT, examinees were asked to complete a questionnaire which contained the following directions:

> Please answer the following questions concerning the Concept Mastery Test that you have just taken. Your answers to these questions will have no effect whatsoever on your Concept Mastery Test score or on your grade in this course.

The questionnaire included the following items:

> 1. Consider those items on the test that you could answer by pure guess only. Approximaely what percentage of these items did you answer?
> 2. Consider those items on the test that you thought you knew, even if you were not quite certain. Approximately what percentage of these items did you answer?

Responses to item 1 were considered to be the reported measure of risk taking (R′); responses to item 2 were regarded as a reported measure of risk taking with partial information (R″).

The study was replicated in five different classes in introductory educational psychology; three classes at the University of California at

Berkeley in the summer of 1965, and two undergraduate classes at the State University of New York at Buffalo during the spring semester of 1966. Females constituted approximately 75% of each class, with the Berkeley classes averaging about 27 years of age compared to about 21 years for the Buffalo groups.

Results

Table 1 presents the sample size, mean, and standard deviation in each class for the disguised (R) and reported (R′) measures of risk taking, and the reported measure of risk taking with partial information (R″). The Berkeley classes are represented by C_1, C_2, C_3; the Buffalo classes are labelled B_1, B_2.

TABLE 1
SAMPLE SIZE, MEAN, AND STANDARD DEVIATION FOR DISGUISED (R) AND REPORTED MEASURE (R′) OF RISK TAKING, AND REPORTED MEASURE OF RISK TAKING WITH PARTIAL INFORMATION (R″)

	C^a_1	C_2	C_3	B_1	B_2
n	26	40	17	51	32
X_R	.81	.75	.80	.72	.54
$X_{R'}$	.43	.40	.47	.28	.17
$X_{R''}$	.84	.88	.83	.71	.60
S_R	.21	.21	.20	.22	.22
$S_{R'}$	.45	.40	.37	.35	.23
$S_{R''}$	.27	.19	.29	.31	.30

[a]C_1, C_2, C_3 represent Berkeley classes; B_1, B_2 represent Buffalo classes.

In each class, the mean disguised risk score was significantly different from the mean reported risk score (.05 level), with the data indicating that the mean of the disguised risk score was higher. In four of the five classes, the reported risk score was more variable than the disguised risk score.

The correlations among the disguised and reported measures of risk taking and the reported measure of risk taking with partial information, are provided in Table 2. Values with an asterisk are significantly greater than 0 at the .01 level for a one-tailed test.

TABLE 2
CORRELATIONS AMONG DISGUISED (R) AND REPORTED (R′) MEASURES OF RISK TAKING, AND REPORTED MEASURE OF RISK TAKING WITH PARTIAL INFORMATION (R″)

	C_1	C_2	C_3	B_1	B_2
$\gamma_{RR'}$	.67*	.70*	.67*	.57*	.44*
$\gamma_{RR''}$	.73*	.38*	.53	.22	.28
$\gamma_{R'R''}$	.46*	.32	.57*	.42*	.28

*Significantly greater than 0 at the .01 level for a one-tailed test.

Discussion

In Table 1 it is seen that in each of the five classes, the subjects displayed higher risk taking (R) than they reported (R′). There are at least three possible explanations for this discrepancy: (*a*) perhaps it is due to the failure of the subjects to recognize the extent of their risk, and hence the subsequent under-estimation of their self-report, (*b*) possibly the subjects recognized the extent of their risk, but for some reason (e.g., social desirability), they systematically reported lower risks than they took. Alternatively, it may be that (*c*) the subjects are accurately reporting their actual degree of risk taking, but for some reason(s) the disguised measure is overestimating the degree of risk. With respect to the last explanation, the assumption that each incorrect response indicates a pure guess provides two plausible reasons for the disguised measure being an overestimate of test-risk. The first reason is based on the effects of partial information, and the second on the effects of misinformation.

With partial information, the subject could not select the best alternative with certainty. Note that if the subject took a risk with partial information, an incorrect response would increase the disguised score (*R*), but the subject presumably would have taken account of the risk in the reported measure with partial information (R″) rather than in the reported measure of risk (R′). Hence, the mean of R would be inflated over that of R′. Indeed, it is interesting to note that in each class, the mean of the reported measure of risk with partial information was approximately the same as the mean disguised risk (see Table 1).

In the case of misinformation, the subject did not need to guess the alternative because he "knew" the answer—unfortunately for the subject, however, the response selected was not the one keyed by the test constructor. Therefore, while the incorrect response thus obtained would increase R, the response would not be considered a risk by the subject, and hence would not enter into either R′ or R″. Once again, therefore, the mean of R would tend to be higher than the mean of R′. Clearly, any of the three previous explanations could account for the observed mean discrepancy between R and R′. However, it would seem that the third explanation is the most plausible.

Table 2 provides evidence that the correlation between the disguised and reported measures of risk is substantial as well as statistically significant. The correlations ranged from .44 to .70, with a median value of .67; the common variance for the two measures therefore ranged from about .19 to .48, with a median value of .45. In examining the multiple correlation between R and both R′ and R″, it was found that no appreciable predictable variance in R was gained by the combining of R″ with R′. Hence it was decided that while risk taking with partial information may

shed some light on the mean discrepancy between R and R′, it provides essentially no further information on the variability of R.

In conclusion, evidence has been offered demonstrating that the correlation between the disguised and reported measure has a median value in the vicinity of .67. Because of this evidence, the construct validity of the disguised measure has been enhanced. While it is clear that the subjects attained higher mean disguised score than they reported, the effects of partial information and misinformation were offered as possible explanations for the observed mean discrepancy.

References

Danielson, L. E. Gambling proneness: Its measurement and expression in examination situations. Unpublished doctoral dissertation, University of Michigan, 1956.

Granich, L. A technique for experimenting on guessing in objective tests. *Journal of Educational Psychology,* 1931, Vol. XXII, 145-156.

Sherriffs, A. C., and Boomer, D. S. Who is penalized by the penalty for guessing? *Journal of Educational Psychology,* 1954, Vol. XLV, 81-90.

Slakter, M. J. The measurement and effect of risk taking on objective examinations. U.S. Department of Health, Education, and Welfare. Final Report: Project No. 5-8428, Contract No. 0E-6-10-239, 1967. (a)

Slakter, M. J. Risk taking on objective examinations. *American Educational Research Journal,* 1967, Vol. IV, 31-43.

Stone, L. A. Reliability of a utility of risk measure. *Psychological Reports,* 1962, Vol. X, 516.

Swineford, F. The measurement of a personality trait. *Journal of Educational Psychology,* 1938, Vol. XXIX, 295-300.

Terman, L. M. *Concept Mastery Test.* New York: Psychological Corp., 1950.

Votaw, D. F. The effect of do-not-guess directions upon the validity of true-false or multiple-choice tests. *Journal of Educational Psychology,* 1936, Vol. XXVII, 698-703.

West, P. V. A critical study of the right minus wrong method. *Journal of Educational Psychology,* 1923, Vol. VIII, 1-9.

Ziller, R. C. A measure of the gambling response-set in objective tests. *Psychometrika,* 1957, Vol. XXII, 289-292. (a)

Ziller, R. C. Vocational choice and utility for risk. *Journal of Counseling Psychology,* 1957, Vol. IV, 61-64.

Alberta Journal of Educational Research
Vol. XIV, No. 2, June, 1968

Expectancy tables were constructed using the grade nine and twelve stanine grades on Provincial Examinations and the comparable rankings on the Scholastic Ability Test of students in a rural Alberta county. The use of the tables for guidance in a small high school is illustrated.

DALE G. ANDERSEN
Washington State University
and
HERBERT E. STILES
Didsbury, Alberta

The Use of Expectancy Tables in Educational Guidance

The rural high school often finds itself at a disadvantage in providing educational guidance services to its students because of the absence of a trained counsellor. Such was the case of the five high schools in a rural Alberta county during the past several years. However, the desire to offer the best that was possible under the circumstances led to a study involving the construction and use of several expectancy tables as guidance tools. This paper reports that study.

The expectancy table is a useful test interpretation instrument and educational guidance device even in the absence of a trained specialist. A Test Service Bulletin (1949) states that:

> The expectancy table is not new; it has been known and used in the test field for more than a quarter of a century. . . . It is merely a grid containing a number of cells [p. 11].

Along the side are categories of standardized test results while along the top are the varying achievement divisions. For each person a tally mark is placed which shows, vertically, his score on a given standardized test, and horizontally, his subsequent achievement in some endeavour.

In the present study expectancy tables were constructed utilizing: (*a*) selected results of the June Examinations of the Examinations Branch, Department of Education, for grade nine students and (*b*) grade twelve status of those same students three years later.

Collection and Reporting of Data

The scores reported in stanines for each grade nine student in the County on the six June Examinations—Reading, Literature, Language, Social Studies, Mathematics, and Science—were collected for the years 1956 through 1961. The percentile rank for each student on the Scholastic Ability Test during the period from 1959 (when they were first administered) through 1961 was also gathered. These data were obtained from either the principals of the schools or from the Department of Education, Edmonton. During the period of the study many students entered County schools from outside the County. Data pertaining to these students were obtained by writing to the principal of the school where the student had written the final grade nine examinations. Few requests were ignored, but the information received was not always complete. Therefore the total number of subjects (N) reported in the tables in this paper varies somewhat.

The grade twelve status of subjects including the results of Grade Twelve Examinations was obtained from the principals of the high schools within the County. The period here was for 1959 through 1964, the period when the original grade nine students had advanced to the twelfth grade.

The first expectancy table constructed, Table 1 below, was for "Total of Stanines." For the six year period of the study the stanines for each grade nine student on the six tests were added. This resulted in actual totals varying from 15 to 54. It was decided to group scores in intervals of five for ease in reporting, thus creating eight possible horizontal rows on the expectancy table.

TABLE 1

EXPECTANCY TABLE FOR TOTAL OF STANINES

Score (Total of Stanines)	N	Grade Twelve Status		
		Matriculation Pass (%)	Diploma Pass (%)	Neither (%)
50-54	19	100	0	0
45-49	57	77	18	5
40-44	99	44	38	18
35-39	121	23	47	30
30-34	178	8	34	58
25-29	125	3	19	78
20-24	74	0	4	96
15-19	14	0	0	100
Total N	687			

Grade twelve achievement entries were made on all expectancy tables according to three categories: (*a*) "Matriculation Pass" denoting the student attained scores of at least 50% on each Grade Twelve Examination and an average of at least 60% on all six Examinations, (*b*) "Diploma

Pass" showing the student attained at least 40% on Grade Twelve Examinations and had accumulated 100 high school credits, and (c) "Neither" which indicated the student did not meet the requirements of either of the first two categories.

After compiling raw totals for each cell on the table these were converted to percentages by row. It was felt this would facilitate interpretation and allow the data to be viewed in a more practical and usable fashion.

In reading Table 1 it is apparent that for students having a total stanine score between 56 and 54 inclusive, 100% of students in the past have attained matriculation. A further illustration is that of students with total scores of stanines between 20 and 24 inclusive in he past none have attained matriculation status, 4 in 100 have gotten a diploma, and 96 in 100 have neither matriculated nor attained a diploma pass.

Table 2 uses the stanine for an individual test, Language. This expectancy table was constructed as before except the row values range only from one to nine since only one test is involved.

TABLE 2

EXPECTANCY TABLE BASED ON LANGUAGE STANINE

Language Stanine Score	N	Grade Twelve Status		
		Matriculation Pass (%)	Diploma Pass (%)	Neither (%)
9	30	96	0	4
8	73	56	31	13
7	97	43	35	22
6	139	17	45	38
5	182	10	28	62
4	112	2	15	83
3	44	5	16	79
2	16	6	6	88
1	4	0	0	100
Total N	697			

The Mathematics stanine was used by itself for constructing Table 3. It would appear that this Examination score is more discriminating than the Language score since the delineations are more clear cut especially at the lower stanine levels.

The last of the expectancy tables (Table 4) that will be presented here involves fewer students. The Scholastic Ability Test (SAT) was administered only throughout the final three years of this study. The standardized test results utilized in this table claimed to reflect ability rather than achievement. An interesting observation in regard to this table is that of those who ranked in row nine of the SAT, the highest rank, only 80% matriculated.

TABLE 3
EXPECTANCY TABLE BASED ON MATHEMATICS STANINES

Mathematics Stanine Score	N	Grade Twelve Status		
		Matriculation Pass (%)	Diploma Pass (%)	Neither (%)
9	29	90	10	0
8	51	62	28	10
7	115	40	40	20
6	156	20	42	38
5	183	11	30	59
4	108	4	13	83
3	46	0	0	100
2	11	0	0	100
1	0	—	—	—
Total N	699			

TABLE 4
EXPECTANCY TABLE BASED ON SCHOLASTIC ABILITY TEST

SAT Score[a]	N	Grade Twelve Status		
		Matriculation Pass (%)	Diploma Pass (%)	Neither (%)
96-99	25	80	4	16
89-95	37	57	35	8
77-88	82	30	35	35
60-76	67	16	36	48
40-59	90	5	32	63
23-39	68	0	16	84
11-22	29	0	7	93
4-10	12	0	0	100
1-3	2	0	0	100
Total N	412			

[a]These are percentile rank scores and were arranged to be roughly equivalent to the stanine intervals used in Table 2 and Table 3.

Guidance Application

The intent of presenting the data shown here to students in the schools of the County is not to tell them what they must do nor to impose decisions on them as to what they should do. Guidance and counselling activities should always be pointed toward human self-determination and the acceptance by the individual of self-responsibility. Frequently, however, students, due to a lack of self-knowledge and inability to assess chances of future success, drop out of school or make choices without any awareness of the probabilities of success or failure involved. They do not realize even in a relative sense their potential to achieve. Some cases of this type would surely be avoided through appropriate presentation of facts.

It is not always the poor student, or the one lacking intelligence who is a dropout. According to French (1966) who conducted a study of dropouts in Ohio schools:

> The greatest loss, . . . is to the individual who is restricted in development and whose contribution of talent to our way of life is thereby reduced. The restriction of self-realization that accompanies withdrawal from school before a student's capability for entering an appropriately high vocational activity is one of the most devastating aspects of the nation's dropout problems [p. 123].

Self-understanding is not the cure-all for the highly complex problem of dropouts, but it may be the answer for many students. It is especially important today because society is asking that students be retained in school for longer periods. The Alberta Legislature, 1966, has decreed that the compulsory school attendance age be raised from 15 to 16 years. This is going to create many problems for the student who is unable to conceptualize himself and his world realistically.

A study by Mueller and Rothney (1960) found that guidance conferences were necessary if superior students were to choose wisely when educational decisions must be made. Choices, after all, are made on the basis of the chooser's prediction of outcomes.

In using the expectancy tables as guidance tools one must be objective. Patterson (1962) states:

> The results of the tests must be allowed to speak for themselves after adequate explanation of the meaning of the scores by the counselor. An illustration of what is meant by objective presentation of the results is the interpretation of a score on a college aptitude test. The statement, "Three out of four persons with scores like yours do not complete the first year of college," is objective. The statement, "Your score indicates that the chances are three out of four that you will not complete the first year of college," while apparently objective, is somewhat loaded. But the statement, "With a score such as yours should not attempt college," is definitely evaluative and judgmental and goes beyond the presentation of information to the client [p. 155].

Since the expectancy tables provided through the present study are from achievement and ability tests, they should be quite reliable, but the person making the interpretation must be objective when using them.

The completed tables may also be used in guidance when dealing with a class or with a group of students. A large copy of each table or an overhead projector transparency could be made. A discussion of each table may be given with explanations and results from previous classes. Students should be encouraged to discuss this information. This would facilitate understanding on the part of each individual as to what the facts are and more important may stimulate individuals to seek out more self-information and information about opportunities. This is especially important in the small secondary school where a guidance counsellor is not on staff.

For some time one of the co-authors of this paper, a high school principal in the County has used the tables presented here to offer as-

sistance to both parents and students. Two illustrations may help in conveying how the tables might be used.

Case One

A boy entering grade twelve with a stanine total of 25 from grade nine examinations was very frustrated. His achievement from the last year was in the low "C" category. He wished to leave school and take technical training in welding and go farming. The parents wanted the boy to complete grade twelve, go to university, and enter a faculty of law.

The mother came to school to talk to the principal. By using the expectancy chart (Table 1), it was possible to show that her son had 3 chances in 100 of matriculating and being able to enter university, and also that her son had 19 chances in 100 of receiving a diploma. The principal also explained that the boy's percentile rank of 48 meant only 37 chances in 100 of receiving a diploma.

The parents' decision after talking to their son, later, was that he could follow his own plan. The boy decided to stay in school and with the reduced parental pressure he had a much more successful year in school.

Case Two

This case involved a boy passing into grade ten in September with a stanine four in mathematics. His percentile rank was 66 on the SAT. The mother wanted the boy to take the academic mathematics.

Using the expectancy table for mathematics stanines she was shown that her son would have 4 chances out of 100 of succeeding in it. However, the mother felt that the general mathematics program was an inferior program and that her son had the ability to take the academic mathematics.

The principal accepted the mother's decision, but felt that the work would be very difficult for the boy as too much background was missing. The boy failed his mathematics in June.

Summary and Implications

This paper has reported the employment of expectancy tables in the educational guidance of rural youth in one county in Alberta. A description of the procedure in constructing these tables based on data from grades nine and twelve was offered and some guidance application reviewed. Two illustrative cases were sketched. The intention of the study was not to work out a device which would give a rigid rule for laying out a student's educational future. It was intended that the expectancy tables would give a realistic appraisal of the situation to the student, parent, or teacher. The greatest value of the tables, then, is as a guidance tool.

The expectancy tables could be refined so that the predictive value would show success of boys and girls separately. The tables could also be modified to be double-expectancy tables[1] combining total stanine scores from grade nine final examinations and the SAT percentiles.

It is hoped that the present paper has presented a procedure and some guidelines that will serve as models for similar projects in other secondary school districts in Alberta. The tables constructed proved useful in the district of their origin.

References

French, J. L. High school dropouts of high ability. *The Vocational Guidance Quarterly*, 1966, Vol. XIV.

Mueller, W. J., and Rothney, J. W. M. Comparison of selected descriptive and predictive statements of superior students, their parents and their teachers. *Personality and Guidance Journal*, 1960, Vol. XXXVIII, 621-625.

Patterson, C. H. *Counseling and guidance in schools.* New York: Harper, 1962.

Expectancy tables—a way of interpreting test validity. *Test Service Bulletin No. 38.* New York: Psychological Corporation, 1949.

---

1 For an explanation of double-expectancy tables see: "Double-entry expectancy tables." *Test Service Bulletin No. 56.* New York: Psychological Corporation, 1966.

Alberta Journal of Educational Research
Vol. XIV, No. 2, June, 1968

Two of the most basic theoretical constructs used by Piaget to explain the acquisition of conservation were identified, described, then tested for their ability to account for the child's performance on a test of conservation. The two constructs were the understanding of transformations and the understanding of states. It was hypothesized that Construct One was more important than Construct Two. This was not found.

DAIYO SAWADA and L. DOYAL NELSON
The University of Alberta

Transformations: Length Conservation

Since the invention by Piaget of using conservation as one important criterion for assessing the degree to which a child has acquired certain fundamental concepts of science and mathematics, much research has been undertaken for the purposes of: (*a*) replicating Piaget's investigations, (*b*) exploring the efficacy of inducing conservation by various training methods, (c) attempting to ascertain the effects of other variables such as age, IQ, type of material, type of property, visual perception, complexity of task, and the like on the acquisition of conservation, and (*d*) criticizing the methodology of Piaget and/or offering constructive alternatives. There has, however, been little experimental examination of some of the theoretical constructs which purport to explain the ontogenesis of conservation in the minds of children. The task of examining some of these theoretical constructs in the context of experimental conditions is of fundamental importance, not only from the viewpoint of child psychology, but also from the viewpoint of the education of children. If we are to teach children to conserve properties of objects or sets of objects, we would hope to base the learning experiences we provide on theoretical constructs which have some experimental operationality as well as hypothetical elegance. Accordingly, this paper (*a*) identifies and describes one of the basic notions thought to underlie the acquisition of conservation, and (*b*) reports a study which attempted to

assess the efficacy of the construct in accounting for the conservation responses of children.

The Theoretical Construct

Since Piaget is the originator of the technique of using conservation as a criterion in the assessment of concept acquisition, and since Piaget has contributed voluminously to the theory of acquisition of conservation in children, a logical source for theoretical constructs is his writing. In selecting a construct of several constructs presented by Piaget, three criteria were kept in mind. The construct had to possess properties such that: (*a*) it could be used to distinguish effectively between nonconservers and conservers, at least in theory, (*b*) it could be used to describe the developmental transition from nonconserver to conserver, and (*c*) it was believed by Piaget to be of fundamental importance. According to Piaget:

> In physical reality there are states, and in addition to these there are transformations which lead from one state to another. . . . In children of the higher stages and in adults . . . any given state is understood to be the result of some transformation and the point of departure for another transformation. But the pre-operational child does not understand transformations. He does not have the operations necessary to understand them so he puts all the emphasis on the static quality of the states. It is because of this, for example, that in the conservation experiments he simply compares the initial state and the final state without being concerned with the transformation (Piaget, 1964, p. 20).

There are two basic constructs given in the above quote: States, and transformations. A nonconserver places all his attentions "on the static quality of the states," while a conserver takes into consideration the transformation which links one state to another. The transition from a nonconserver to a conserver is described in terms of an understanding of transformations. Thus, the notions of state and transformations satisfy the selected criteria.

States and Transformations: A Rationale

The idea of using states and transformations as a theoretical construct designed to render some structure to our thought is an idea which has been used in many different disciplines. Langer, for example, in her book on logic (1953, p. 20) uses the notions of states, and transformation from one state into another, to distinguish between our knowledge *of* something, and our knowledge *about* something. The former is a direct sensuous knowledge of things, the kind of knowledge a baby has of the crib, of the blankets, or of its mother's breast. However, the baby does not know *about* his mother's breast. He does not know about the relative location in space, although at times he will know of their location;

he does not know how various cultures view them, although he may know of his own reaction to them; he does not know about the effects of lactation, although he certainly may know of these effects and so on. In other words, to have knowledge *about* an object (or a state) requires more than just direct sensuous contact with it. A young school boy learning linear measurement by rote may have knowledge *of* measurement since he has had direct sensuous contact with rulers and other measuring apparatus, but he may have no knowledge *about* measurement, and indeed he may not even conserve length (Lovell, Healey, and Roland, 1962). To have knowledge *about* linear measurement is indeed to go far beyond the mere application of a measuring stick. Knowledge *about* linear measurement (knowing for example that length is additive, or that rotations and translations do not change length) gives meaning to knowledge *of* linear measurement (i.e., gives meaning to the mechanical application of a measuring stick).

In summary, the understanding of transformations which relate one state to another turns knowledge *of* each state into knowledge *about* the states. Operational thought could be thought of as emerging when a sufficient understanding of transformations enables the child to integrate these states into a reversible system of states, this systematization emerging out of a set of transformations linking the states, each set of transformations inducing the inference of invariants (or properties) which characterize the states being systematized, and hence give rise to conservation (and many other refinements).

For example, when a sufficient understanding of the ideas embodied in euclidean transformations enables the child to systematize space into a frame of reference or container for physical movement, then each location of a solid body becomes reversibly related with any other location assumable by that body. The set of euclidean transformations, therefore, defines certain invariants such as length, area, or shape which emerge as properties characterizing the objects and, thus, sets the basis for the conservation of these properties. Likewise, when a sufficient grasp of the basic ideas embodied in topological transformations enables the child to conceive of each state of a piece of plasticine as being merely a different form of the same basic substance, each form being transformable into any other such form, then the inducement of the inference that some basic underlying thing is remaining invariant under all such transformations sets the basis for the emergence of the conservation of substance, weight, and other properties which remain invariant.

In summary, it is the growth in knowledge *about* states, or, in other words, the growth in knowledge *of* the *transformations* which relate one state to another state, which transforms a nonconserver into a conserver.

As an interesting aside, it might be conjectured that it is knowledge *about* transformations which leads the child from concrete operations to formal operations.

Experimental Studies

The studies which have the most direct bearing on the problem of identifying important variables thought to underlie the acquisition of conservation are those which attempted to ascertain the efficacy of various training methods for inducing conservation in children. Of the several studies done, it is probably safe to describe only three (Beilin, 1965; Wallach and Sprott, 1964; and Sigel, Roeper and Hooper, 1966) as having experienced, to a significant degree, success in inducing conservation. However, before considering these studies, it may be pertinent to discuss a study by Greco (1959, cited in Smedslund, 1961a, and in Lauredeau and Pinard, 1962) who tested the efficacy of empirical rule learning acquired by external reinforcement, and interpreted his findings in terms of the understanding of transformations.

Greco's apparatus consisted essentially of three coloured balls attached to a rod which was pulled through a cardboard tube which could be rotated in multiples of 180 degrees. The child observed the balls as they entered the tube, and after one or two rotations, tried to predict the colour of the first ball exiting. Two groups, designated D and A, of children aged 5-10 to 7-1 were given training; group D (the empirical group) practiced the outcomes of one and two rotations *separately* to perfection, while group A learned a *mixture* of one and two rotations to perfection. A posttest given some one to three months later revealed that nearly all of group D had forgotten everything, while group A retained almost everything; furthermore, group A showed considerable but not complete generalization to "n" rotations. Greco concluded that learning occurred in two ways: by learning the outcomes of one and two rotations separately as empirical rules which have no "necessary" quality about them, and by profound internal re-structuring forced by the more complex nature of the presentation of a mixture of rotations. Greco inferred that the mixture led the subjects to search for relations between one and two rotations, and to understand that the composite transformation of two rotations is simply the product of one rotation followed by another. Thus, Greco's explanation highlights the role of the understanding of transformations (rotations in this case).

Wallach and Sprott (1964) also interpreted the success of their training method in terms of an understanding of transformations. Using a training method of inducing number conservation which (*a*) required the child to predict whether a one-to-one correspondence could be re-established between two sets of discrete complimentary objects (dolls

in beds) after the objects had undergone an additive-subtractive or spatial transformation or both, and (*b*) required the child to attempt to reconstruct the one-to-one correspondence, the investigators found that the method was successful beyond the .001 level of significance. More conclusively, a posttest given from two to three weeks later showed that the experimental group maintained their conservation. Wallach and Sprott reasoned that the crucial variable underlying conservation is "reversibility," "reversibility" being definable in terms of transformations as the following summary quotation indicates:

> a property is regarded as conserved under an operation removing a defining attribute, when it is expected that the inverse operation will lead to the attribute's reappearance at the same value [p. 1069].

It should be noted that the term "operations" is used in the above context precisely the way the present writer uses the term "transformation," and parenthetically, from a mathematical point of view, any particular operation can be conceived of as being a special kind of transformation, and Piaget's theories are extensively rooted in mathematics.

However, Lovell and Ogilvie (1961), using the now classic plasticine balls experiment to test for the conservation of weight and substance, reported that 46% of the nonconservers showed evidence of being able to recall that the balls were equal before the shape transformation was applied to one of them. Defining reversibility as "the capacity to show awareness of the equality of the balls before the experiment (p. 139)," the authors concluded that reversibility is not a sufficient condition for conservation. Supposing that the discrepancies between Wallach and Sprott's conclusion and that of Lovell and Ogilvie is not accounted for by the difference between number conservation and weight conservation, then what is the explanation of these contrary findings? Wallach and Sprott suggest that mere "dispositional reversibility" (such as defined by Lovell and Ogilvie) does not require the child to actually imagine the inverse transformation which will re-establish the equality. In the terminology of the present paper, the criterion of reversibility in Lovell and Ogilvie's definition is a *state* criterion; the child has only to recall the equality state prior to the transformation. The criterion in Wallach and Sprott's definition is a dynamic criterion; the child must actually think of the *transformation* (or operation) which will re-instate the equality state. What Wallach and Sprott are suggesting is that a child will conserve if he realizes that every spatial transformation, indeed, any conserving transformation, necessarily implies its own inverse. They point out that conservation is not necessarily deducible from the fact that every shape transformation has an inverse, but that to *assume* that this is the case makes for cognitive economy.

Reversibility also played a part in a training study done by Sigel, Roeper, and Hooper (1966). However, the reversibility training in this

study was only one of three types of training given to all the children, the other types being "multiple classification" and "multiple relationality." It is impossible to determine which of these three aspects contributed most to the success of the training, or whether the interaction of these aspects was the crucial variable.

Another recent study in which some success was experienced in inducing number conservation in childern was Beilin's (1965). In general, his procedure consisted of the familiar pretest-training-posttest design which he used on 170 children aged 5 to 11. The only training method of the four used which produced significantly greater gains than those shown by the control group was a highly verbal method Beilin called "Verbal Rule Instruction." It appears highly unlikely that "Verbal Rule Instruction" could induce any profound cognitive reorganization, and perhaps profound cognitive reorganization is not required. This, however, would contradict the findings of Piaget (1950, 1960), Wohlwill and Lowe (1962), Smedslund (1961a, b, c, d, e, f), Feigenbaum and Sulkin (1964), and Ervin (1960). Perhaps the explanation is that Beilin's posttest apparatus, materials, and procedure were almost identical to those used in the training sessions so that superficial cognitive reorganization verbally induced would be sufficient to give correct responses on the posttest. Such superficial learning would not be expected to transfer to more different contexts involving the same conservation principles, and Beilin (1965) and Beilin, Kagan, and Rabinowitz (1966) have found that indeed it does not.

In summary, it may be concluded that an understanding of transformations has played a rather important if not crucial role in the training methods that were successful in inducing conservation. It appears, therefore, that the role of transformations in the acquisition of conservation merits a closer examination. It may even be hypothesized that the most crucial variable relating to the acquisition of conservation is an understanding of transformations.

Method

In order to test the hypothesis that an understanding of transformations is an important factor in the acquisition of conservation, the investigator translated this general problem into a specific one dealing with the conservation of length. A special test of length conservation consisting of 32 items with built in euclidean transformational and state properties classified in certain basic ways was constructed on the assumption that, when such a test was factor analyzed, the hypothesis of the importance of transformations could be answered by ascertaining whether the underlying dimensions in the test could be best interpreted in terms of transformations or not.

Subjects

The sample consisted of 62 kindergarten and grade one pupils aged 5-4 to 8-0 randomly selected from a population of 196 such pupils attending schools servicing the Canadian Forces Base at Griesbach near Edmonton, Alberta. The vast majority of the fathers of the children were of the Corporal or lower rank.

Procedure

For this study, an essentially nonverbal method was used in testing for the conservation of length under the assumption that a nonverbal method would give a "purer" measure of conservation (Braine, 1959, 1964; Sawada and Nelson, 1968).

To remove the dependence on the word "length" to communicate the concept of length to the child, calipers were used to define to the child that attribute about which he was to make decisions.[1]

A training session of about 12 to 15 minutes per S was given to teach the child how to use the calipers in conjunction with a three-choice response apparatus in order to facilitate S's decision-making about how a given transformation affected the length of various rods. The child was first taught to arrange the rods in a straight compact line before attempting to apply the calipers of which there were three sizes—9 cm., 9.5 cm., and 10 cm. It was arranged so that the calipers always fitted the rods in one of only three ways:

1. The calipers were .5 cm. too short for the calipers.
2. The calipers were just right for the rods.
3. The calipers were .5 cm. too long for the calipers.

Each of the three different "fits" described above was associated with one of three "doors" on the response apparatus, the association being facilitated by placing a "model fit" above each of the doors and by training S to apply a pair of calipers to a set of rods, to note how the calipers fit, to locate the equivalent model fit, and finally, to open the corresponding door which revealed a candy reward. Most S looked upon the above procedure as a game, and all 62 S could easily find the candy. The "game" was slightly modified by having S predict the fit after the calipers were removed and the rods given a transformation. The resulting procedure was as follows:

1. A rod or set of rods and a pair of calipers were placed in front of S.
2. S applied the calipers to the rods noting the fit.
3. E removed the calipers and applied a transformation to the rods.
4. S predicted how the calipers would fit if he were allowed to put them on again. "How would they fit *if* you tried to put them on again? Like this, or this, or this?" E asked, pointing at the model fits. "Find the candy."

---

1 A fuller description of the testing technique used in the present study is given in Sawada and Nelson (1968).

Three practice items were given using steps 1 to 4 above, and all S appeared to understand what was required. The test was then given using steps 1 to 4 for each of the 32 items. A fuller description of the testing technique exists elsewhere (Sawada and Nelson, 1968).

The Test

Since the important variables under study are transformations, and states, a variety of euclidean transformations and a variety of state attributes were incorporated into the 32 items resulting in items which could be classified into several subsets according to the kind of transformation or state attribute embodied in the item. The transformations could be classified into the following kinds:

1. Translations—e.g., a wooden rod is pushed 20 cm. horizontally.
2. Rotations—e.g., a wooden rod is rotated 60°.
3. In-plane transformations—a transformation which can be executed entirely within one or two dimensions.
4. Out-plane transformations—a transformation which requires three dimensions before it can be executed: i.e., the plane must be conceived of as immersed in a three-space.

The above 4 kinds of transformations all leave length invariant. The test would be crucially incomplete without transformations which alter length. Therefore, a fifth classification was included.

5. Noninvariant transformations—a transformation which alters the length of an object.

It should be noted that for transformations 1 to 4 above, a precise mathematical rationale can be given in terms of orthogonal line reflections (see Sawada, 1966, pp. 122-25).

The state attributes were of the following kinds:

1. Multiple segmented lengths—a length exhibited by an object consisting of two or more discrete parts.
2. Single segmented length—a length exhibited by an object consisting of a continuous whole.
3. Relations—relations of less than, equal to, and greater than (F, =, J) between the rod and calipers.
4. Material—wood, plasticine, cigarette.
5. Illusions—these were variants of (*a*) Muller-Lyer, (*b*) Baldwin, (*c*) Opel inverted T, (*d*) the use of two perspective lines. These illusions were used in an attempt to make the rods appear to have increased in length.

Table 1 gives a summary of the above dimensions, and lists the test items incorporating each of the dimensions.

Hypothesis and Data Analysis

The hypothesis can be explicitly stated as follows: When a factor analysis is performed on the conservation test used in this study, the dimensions which will define the factors will be transformational dimensions rather than state dimensions.

TABLE 1

DIMENSIONS BY WHICH THE TEST ITEMS ARE CLASSIFIED

Dimensions	Items measuring the dimension
Translations	1, 2, 3, 4, 5, 6, 7, 8,
Rotations	9, 10, 11, 12, 13, 14, 15, 16,
Noninvariant	17, 18, 19, 20, 21, 22, 23, 24,
In-plane	2, 3, 5, 8, 9, 12, 14, 15,
Out-plane	1, 4, 6, 7, 10, 11, 13, 16,
Multiple Segments	3, 4, 7, 8, 11, 12, 15, 16,
Single Segments	1, 2, 5, 6, 9, 10, 13, 14,
$<$ (less than)	5, 6, 11, 12, 13,
$=$ (equal to)	1, 2, 3, 4, 9, 10,
$>$ (more than)	7, 8, 14, 15, 16,
Wooden rods	17, 18,
Plasticine rods	21, 22,
Cigarettes	23, 24,
Illusions	25, 26, 27, 28, 29, 30, 31, 32,
I ($=$)	25, 26, 27, 28,
I ($\neq$)	29, 30, 31, 32,

The study was designed as a factor analytic study rather than a training study because the investigator believed that it is not known just what transformational variables are the crucial ones. Instead of incorporating one or two transformational variables in a training method, the investigator included several transformational variables, systematically classified in some way, in a test of conservation, and performed a factor analysis on the test to ascertain whether understanding of transformations is indeed an important factor, and, if so, which transformations are important.

The Principal Axis method of factoring programmed for the IBM 7040 was employed using the fortran coded programs of the Division of Educational Research Services of The University of Alberta. Since the Principle Axis method extracts factors in the order of the amount of the variance accounted for, it was assumed that any dimension crucial to the performance on the test would show itself in at least one of the first few factors. Accordingly five factors were extracted and interpreted.

In order to facilitate the interpretation of the factors, the factor matrix was rotated to approximate "simple structure" using the Equimax criterion programmed for the IBM 7040. Table 2 presents the rotated matrix of factor coefficients, and Table 3 presents the transformation matrix linking the rotated matrix to the unrotated matrix.

Interpretation of the Factors

For the purposes of facilitating the identification of the factors, Table 4 was drawn up denoting the major configurations in the factor pattern. The table is based on factor coefficients, which are at least as high as .350; coefficients between .350 and .499 are designated with an "m" for "medium" factor loading, and coefficients between .500 and 1.000 are designated by an "H" for "high" factor loading.

Items 1 to 16 load on factors I and II while items 17 to 32 load on factors III, IV, and V. The column headed by "Relevant Dimension" aids in the interpretation of the factors. Factor III is essentially an I(=) factor (illusions involving equality between the rod and the calipers).

TABLE 2

ROTATED FACTOR MATRIX

Test Item	Factor I	II	III	IV	V	Commu-nality
1	.390	.417	.191	—.124	.271	.451
2	.363	.738	—.034	—.098	.177	.719
3	.698	.468	.234	—.044	.169	.791
4	.646	.540	.218	.069	.106	.776
5	.041	.643	.297	.243	—.085	.557
6	—.083	.808	.030	.154	—.088	.691
7	.462	.398	.325	.125	—.270	.566
8	.642	.071	.061	.052	.053	.427
9	.512	.638	—.012	—.013	—.070	.675
10	.561	.461	.062	—.170	.250	.623
11	.591	.232	—.044	—.118	.317	.519
12	.804	.075	—.047	.164	—.107	.692
13	.595	—.053	.452	—.035	.120	.572
14	.139	.714	.100	—.125	—.031	.556
15	.735	.400	.118	.064	—.101	.728
16	.671	.194	—.036	—.008	—.083	.496
17	—.296	.104	.212	.225	.483	.427
18	—.087	.005	.024	.339	.197	.162
19	—.017	.130	.140	.611	.071	.415
20	.207	.182	—.186	.307	.414	.377
21	—.059	—.030	—.125	.190	.759	.632
22	—.100	—.130	—.331	.388	—.130	.304
23	.027	—.221	.164	.309	.352	.296
24	.100	—.082	.294	—.072	.675	.564
25	—.054	.073	.705	.147	—.003	.527
26	.233	—.138	.551	—.124	.351	.515
27	—.035	.114	.752	—.001	.186	.614
28	.142	.180	.730	—.034	—.088	.594
29	—.006	—.065	.262	.494	—.015	.318
30	.212	.222	—.142	.402	.111	.289
31	.078	—.042	—.065	.755	.068	.587
32	—.059	—.011	—.131	.659	.111	.467
Variance accounted for	5.005	4.095	2.923	2.596	2.259	16.927
% of Common Variance	29.86	24.19	17.27	15.34	13.34	100.00

TABLE 3

TRANSFORMATION MATRIX

—.735	—.613	—.268	—.032	—.108
—.182	—.136	.503	.593	.587
—.129	—.122	.706	—.686	.019
—.550	.624	.236	.222	—.452
—.329	.449	—.349	—.359	.663

Note: If T is the transformation matrix and if U is the unrotated matrix and R the rotated matrix, then UT = R.

TABLE 4

'SCHEMATIC' FACTOR MATRIX

Test item	Relevant Dimensions		Factor I	II	III	IV	V	Test item
1	Single,	=	m*	m				1
2	Single,	=	m	H				2
3	Multi,	=	H	m				3
4	Multi,	=	H	H				4
5	Single,	<		H				5
6	Single,	<		H				6
7	Multi,	>	m	m				7
8	Multi,	>	H					8
9	Single,	=	H	H				9
10	Single,	=	H	m				10
11	Multi,	<	H					11
12	Multi,	<	H					12
13	Single,	<	H		m			13
14	Single,	>		H				14
15	Multi,	>	H	m				15
16	Multi,	>	H					16
17	Non-inv						m	17
18	Non-inv							18
19	Non-inv					H		19
20	Non-inv						m	20
21	Non-inv						H	21
22	Non-inv					m		22
23	Non-inv						m	23
24	Non-inv						H	24
25	I(=)				H			25
26	I(=)				H			26
27	I(=)				H			27
28	I(=)				H			28
29	I(≠)					m		29
30	I(≠)					m		30
31	I(≠)					H		31
32	I(≠)					H		32

*"m" designates a "medium" factor loading (.350 - .499)

"H" designates a "high" factor loading (.500 - 1.000)

Factor V is essentially a noninvariant transformation factor. However, the identity of factors I and II is by no means obvious. By using the dimensions single and multiple segmented lengths, and the relations $<$, $=$, $>$, some identity can be given to factors I and II. The primary dimensions seem to be single and multiple segmented lengths, with factor I being a multiple segment factor, and factor II being a single segment factor. However, this dichotomy is modified by the relations $<$, $=$ $>$. When the relations is $<$, the items tend to load higher on factor I. When the relations is $>$, the items tend to load higher on factor II. When the relations is $=$, the items tend to equalize on both factors. These remarks on the interpretation of factors I and II can be verified in general by examining Table 4.

For the purposes of summary, it should be noted that all items loading on factors I and II are items which leave length invariant. Thus, if factors I and II are grouped together, they can be thought of as embodying the notion of "invariant transformations" and together can be contrasted with factor IV which may be called "noninvariant transformations." Hence, the major transformational variable in the test can be represented by the classification of transformations into those which leave length invariant, and those which do not. A finer classification of transformation does not seem justified on the basis of the above analysis. The state variables I($=$) and I($\neq$), defining factors III and IV respectively, emerge as the major state variables; single segments, multiple segments and the $<$, $=$, $>$ relations emerge as secondary state variables, defining as they do, the distinction between factors I and II.

Discussion

The results indicate that understanding transformations is of no more importance than the understanding of states as concerns the acquisition of the conservation of length. The fact that state dimensions contributed as much to the identification of the factors as did the transformational dimensions suggests that training studies designed to teach children to acquire conservation should incorporate in their training sessions mental operations which explicitly incorporate both state and transformational dimensions. Moreover, transformations which change the magnitude of the attribute must be included as well as transformations which leave the magnitude of the attribute invariant.

The Sigel, Roeper, Hooper study (1966), which was successful in training children in acquiring conservation, used a training method which was designed to increase the child's awareness and facility with the mental operations associated with: multiple classification, multiple relationality, and reversibility. Multiple classification and multiple relationality (as described by Sigel, Roeper, and Hooper) are dependent essentially

on state attributes whereas reversibility is dependent essentially on understanding physical transformations. It appears then that Sigel and his colleagues included both aspects in their training. Had they left out reversibility, their training method may not have been nearly as successful as it was.

This study concludes with the inference that the understanding of state attributes and of transformations from one magnitude of the attribute to another are equally important to the acquisition of conservation.

Moreover, this inference appears likely from a theoretical point of view. Just as transformations give meaning to the states transformed, knowledge *of* state attributes gives meaning to the transformations applied. Without direct knowledge *of* states, the effect of the transformation on the states becomes indeterminant. With meaningless states, there can only be meaningless transformations. With meaningless transformations there can only be rote knowledge *of* states. As Piaget might say, first the child must differentiate reality into attributes of states, and transformations from one magnitude of the attribute into another; then he must co-ordinate these in an operational system of states and transformations.

References

Beilin, H. Learning and operational convergence in logical thought development. *Journal of Experimental Psychology,* 1965, Vol. 2, 317-339.

Beilin, H., Kagan, J., and Rabinowitz, R. Effects of verbal and perceptual training of water level representation. *Child Development,* 1966, Vol. 37, 317-329.

Braine, M. D. S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. *Psychological Monographs,* 1959, Vol. 73, (Whole No. 475).

Braine, M. D. S. Development of a grasp of transitivity of length: A reply to Smedslund. *Child Development,* 1964, Vol. 35, 799-810.

Ervin, S. M. Training and logical operation by children. *Child Development,* 1960, Vol. 31, 555-564.

Feigenbaum, K. E., and Sulkin, H. Piaget's problem of conservation of discontinuous quantities: A teaching experience. *Journal of Genetic Psychology,* 1964, Vol. 105, 91-97.

Langer, S. M. *An Introduction to symbolic logic.* New York: Dover, 1962.

Laurendeau, M., and Pinard, A. *Casual thinking in the child.* New York: International Universities Press, 1962.

Lovell, K., Healey, D., and Roland, A. D. The growth of some geometrical concepts. *Child Development,* 1962, Vol. 30, 751-767.

Lovell, K., and Ogilvie, E. A study of the conservation of substance in the junior school child. *British Journal of Educational Psychology,* 1960, Vol. 30, 109-118.

Piaget, J. *The psychology of intelligence.* London: Routledge & Kegan, Paul, 1950.

Piaget, J. *The child's conception of geometry.* New York: Harper & Row, 1960.

Piaget, J. Development and learning. In R. Ripple and V. Rockcastle (Eds.), *Piaget rediscovered: A report of the conference on cognitive studies and curriculum development.* Cornell University, 1964. Pp. 7-20 (mimeo).

Sawada, D. Transformations and concept attainment: A study of length conservation in children. Unpublished Masters Thesis, The University of Alberta, 1966.

Sawada, D., and Nelson, L. D. Conservation of length: Methodological considerations. *Alberta Journal of Educational Research,* 1968, Vol. XIV, No. 1.

Sigel, I. E., Roeper, A., and Hooper, F. H. A training procedure for acquisition of Piaget's conservation of quantity: A pilot study and its replication. *British Journal of Educational Psychology,* 1966, Vol. 36, 301-311.

Smedslund, J. The acquisition of conservation of substance and weight in children: I. Introduction. *Scandinavian Journal of Psychology,* 1961, Vol. 2, 11-20. (a)

Smedslund, J. The acquisition of conservation of substance and weight in children: II. External reinforcement of conservation of weight and of the operations of addition and subtraction. *Scandinavian Journal of Psychology,* 1961, Vol. 2, 71-84. (b)

Smedslund, J. The acquisition of conservation of substance and weight in children: III. Extinction of conservation of weight acquired "normally" and by means of empirical controls on a balance scale. *Scandinavian Journal of Psychology,* 1961, Vol. 2, 85-87. (c)

Smedslund, J. The acquisition of conservation of substance and weight in children: IV. An attempt at extinction of the visual components of the weight concept. *Scandinavian Journal of Psychology,* 1961, Vol. 1, 153-155. (d)

Smedslund, J. The acquisition of conservation of substance and weight in children: V. Practice in conflict situations without external reinforcement. *Scandinavian Journal of Psychology,* 1961, Vol. 2, 156-160. (e)

Smedslund, J. The acquisition of conservation of substance and weight in children: VI. Practice on continuous versus discontinuous material in conflict-situations without external reinforcement. *Scandinavian Journal of Psychology,* 1961, Vol. 2, 203-210. (f)

Wallach, and Sprott, R. Inducing number conservation in children. *Child Development,* 1964, Vol. 34, 1057-1071.

Wohlwill, J. F., and Lowe, R. C. Experimental analysis of the development of conservation of number. *Child Development,* 1962, Vol. 33, 153-167.

BOOK REVIEW

THE SPECIFICATION AND MEASUREMENT OF LEARNING OUTCOMES

By: David A. Payne (Waltham, Massachusetts: Blaisdell, 1968).

The author states that the *Specification and Measurement of Learning Outcomes* is intended for:

1. Classroom teachers,
2. Classes in test construction at both undergraduate and graduate levels,
3. Education and psychology classes requiring an overview of measurement.

The ten chapters of the book may be divided into four broad sections. The first section, chapters one to five, after a brief orientation, deals with test planning and construction. The interrelationship of educational objectives, teaching, and measurement-evaluation activities is brought out. Next, procedures for developing a test are presented, beginning with the table of specifications and then leading into types of instruments, item writing, and the use of essay tests.

The second section, chapters six to eight, deals with deriving and applying test data. The section begins with statistical procedures, then proceeds to test validity and reliability, and finally, to item analysis by hand methods.

The third and fourth sections consisting of one chapter each deal respectively with the use of standardized achievement tests and assigning grades. The approach used by the author is one of dealing with the problems commonly found in the two areas. Such problems as choosing an achievement test, sources of pertinent information, misconceptions regarding achievement tests, are discussed in chapter nine. In the final chapter, the more important problems discussed are: the relative or absolute standard, achievement and effort as bases for marks, combining scores, and adjusting grades.

The book suffers from over-summarizing of complex material. This is to be expected in view of the fact that the author tried to cover in about 200 pages most of what is found in much longer measurement texts plus other sources such as Bloom's and Krathwohl's "Taxonomies." Tabular summaries of such things as "principles of oral examinations" and "critical questions useful in reviewing educational achievement test items" do not seem to promote learning. Also, the chapter on statistical

procedures seems to expect too much of the reader if he is, indeed, part of the audience for whom the book is intended. A student or teacher with no knowledge of statistics, instead of acquiring some knowledge, will likely fall victim to "symbol shock."

The chapter dealing with item analysis was disappointing to this reviewer. The "show of hands" technique is recommended for small groups of pupils and a brief description is given. However, there is no suggestion that machine-based methods are becoming increasingly available to teachers. This seems to be an oversight, in that teachers more commonly deal with several sections of a course and the groups are therefore too large to be involved in the simple hand methods. It must be conceded that this viewpoint represents a personal bias.

It is possible that the book would be of value to classroom teachers seeking to gain a quick review of the field of educational measurement, especially if it were used in a workshop situation. Students in courses which have a unit on measurement might also find the book of value; however, it is likely that excepts from well known texts such as Ebel's, Thorndike and Hagen's, or Downie's would be more valuable. The book would not be of much value as a text or primary reference for undergraduate or graduate measurement courses.

It is obvious that the author delved deeply into the writings on educational measurement. The suggested readings found at the end of each chapter, and the list of references at the end of the book, provide a valuable summary of the literature in the field.

V. R. Nyberg
The University of Alberta

NOTES ON CONTRIBUTORS

DALE ANDERSEN is an Assistant Professor of Education at Washington State University, Pullman, Washington.

RAM K. GUPTA is an Assistant Professor in the Department of Educational Psychology, The University of Alberta. He holds B.Sc., L.T., and M.Ed. degrees from India; and M.A. from London, and a Ph.D. from Minnesota. He has been employed on a number of research projects in Minnesota and has published in a number of journals. His special interests are in measurement theory, research design, and statistics.

H. D. HEMPHILL, B.A., M.Ed., is a graduate of the Universities of British Columbia and Alberta. He is currently completing doctoral requirements at The University of Alberta, Department of Educational Administration. He is the author of articles on the teaching of modern mathematics, educational finance and the educational use of automatic data processing.

R. S. MACARTHUR, whose doctoral study was carried out at the University of London, has for some years been a Professor of Educational Psychology at The University of Alberta. His current research interests are on the theme of cognitive development in cross-cultural context. His numerous publications include reports of work with native peoples of the Canadian North and of Central Africa.

L. D. NELSON obtained his Ph.D. from the University of Minnesota after graduating from The University of Alberta. He is co-editor of MATHEX published by Encyclopedia Britannica of Canada. He is a member of the Publications Committee and a member of the Editorial Board for a yearbook on the history of mathematics, National Council Teachers of Mathematics.

VERNER R. NYBERG is an Associate Professor at The University of Alberta. He holds B.Sc., B.Ed., and M.Ed. degrees from The University of Alberta and an Ed.D. from U.C.L.A. His special interests are psychological measurement and construction of classroom tests.

AUBREY H. RODEN studied at The University of California, Berkeley and at The University of Alberta, Canada. He has published articles in journals such as the *British Journal of Psychology.*

DAIYO SAWADA holds a joint appointment as an Assistant Professor in Elementary Education and in the Division of Educational Research Services at The University of Alberta. He has had articles published in the *Mathematics Teacher* and the *Alberta Journal of Educational Research.*

MALCOLM J. SLAKTER studied at Syracuse University and at The University of California, Berkeley. He has had articles published in journals such as *Biometrika* and the *American Educational Research Journal.*

HERBERT E. STILES received his Master of Education degree from The University of Montana. He is now the principal of Didsbury High School, Didsbury, Alberta.

The Alberta Journal of Educational Research

Vol. XIV, No. 3 *September, 1968*

CONTENTS

BOOK REVIEWS

COMMITTEE ON EDUCATIONAL RESEARCH
The University of Alberta

A SPECIAL ISSUE
THE USE OF COMPUTERS IN EDUCATION

A special issue on The Use of Computers in Education is being planned for 1969. Several manuscripts are in hand but scholars interested in this area are invited to submit manuscripts or give notice of intention to do so.

Editorial Comment

In the June issue of the Journal, an invitation was extended to interested scholars to submit articles for a special "theme" issue which is being planned for 1969. The theme for the special issue is to be *The Use of Computers in Education.* We would appreciate being advised of intention to submit an article so that tentative plans for the publication may be made as soon as possible.

As suggested in the previous editorial comment, manuscripts of two types are pertinent to the theme of the issue. One type of article will set forth a rationale for the "involvement" of computers in education. These articles will be broadly based to set the stage for an insightful look at the role of computers in education. The second type of paper will be more specific, dealing with the use of computers in research, instruction, and curriculum building in education. A statement of the contents of the article you propose to submit would be useful to the Editor as plans for the issue are formulated.

* * *

In this issue of the Journal, we have three basic types of research reported in the several articles. One report deals with an evaluation of an approach to historical enquiry. Two articles report the results of experimental studies. The third type of article might be classified as an investigation—an investigation designed to explore the relationships of variables operating in a specific situation.

In the first article, Hodysh discusses the strengths and weaknesses of Berlin's approach to historical enquiry. On the basis of his evaluation, he suggests the modification of the research method that would be necessary in a consideration of the complexities of educational history.

The two experimental studies reported are those of Knowles and Boersma, and of Briggs. Knowles and Boersma report a study of a narrow, well-defined element in the process of concept formation. They compare groups of culturally different children on reversal-nonreversal tasks designed to reveal the development of mediational responses. Briggs presents the results of a preliminary study comparing the results of two counselling methods, one of which is a relatively untried method of directive group counselling. His study, though it does not support definitive statements of the superiority of one or the other methods, does offer sufficient evidence to justify further experimentation with directive group counselling.

In their report of another facet of a study of the merit of failing grades, Storey and Patterson present findings concerning two groups of high school students. One group of students had repeated a high

school social studies course and the other group, similar in achievement, had not repeated the course. Their findings do not support failure as a means to preparing students better for further study in a subject area.

Conklin and Ogston present the results of an investigation of the relationship of some measures used to select students for university entrance and the grades attained by these students after one year of university study. Their results point up some serious weaknesses in many often-used measures. The predictive value of measures other than high school average proved to be low.

MacLeod and Knill extend the investigations of the adolescent subculture of the high school. Specifically, they report the relationships of social class to participation in students' councils in the high school. Their results lead them to question whether students' councils actually provide training in democratic processes.

* * *

The response of subscribers to the choice of a subscription year—January to December or September to June—has been positive. We are pleased that this modification in our business procedures has been well-accepted.

P.A.L.

Alberta Journal of Educational Research
Vol. XIV, No. 3, September, 1968

According to Isaiah Berlin, the complexity of historical events prevents the study of history from becoming "scientific." The author presents the view that Berlin's theory limits the techniques of historical inquiry. The historian of education, in keeping with the principles of historiography, should select any procedures which enable him to interpret the past.

HENRY W. HODYSH
The University of Alberta

Historical Inquiry and Educational History

In his polemic "The Concept of Scientific History" Isaiah Berlin (1961, pp. 1-31) questions the "scientific" approach to historical inquiry. The complexity of historical events is fundamental to his position. Although his ideas are applicable to the history of education, much of what he says is open to question.

Berlin's arguments are presented through a comparison of the natural and social sciences of history.[1] After indicating variant trends in the approach to historical study, Berlin associates the inductive-deductive principles of science with the view that historical events move on the "river of time," and that their never-ending pattern must exhibit regularities capable of being condensed into laws. Following the refutation of this position, Berlin states what he considers to be a crucial difference between natural science and history.

> It seems clear that whereas in history we more often than not attach greater credence to particular facts than to general propositions, however well supported, from which these facts could in theory be deduced, in a natural science the opposite seems to be the case: there it is more rational to place credence in a properly supported general theory . . . than in specific facts [p. 9].

He emphasized the generalizing nature and strict logical interconnectability of laws in the natural sciences, and loose, general statements of history whose credulity must be supported by "judgement" and "intuition." Although historians may attempt to write a scientific history,

[1] Depending on the context, history might be considered as either that which has actually happened or that which has been recorded. See W. Dray, Philosophy of history. (Englewood Cliffs, N.J.: Prentice-Hall Inc., 1964. P. 1.)

the outcome is necessarily vague, inconclusive, and impracticable because of the mass of unsifted, interdependent factors that bear upon any historical question. Berlin's attack is continued upon the speculations of Hagel and Toynbee and specialized disciplines such as the history of technology. He admits that these systems satisfy some of the criteria of natural scientists, "but only at the expense of leaving out the greater part of what is known in the life of human beings whose histories are in this way recorded [p. 13].

The effectiveness of scientific models, which are directly dependent on their simplicity, is the next topic of Berlin's interest. Since history is infinitely complex, the scientific model, or any model for that matter, is of doubtful if any value. He writes:

> . . . the greater the number of similarities we are able to collect . . . the more successfully we abstract, the simpler our model will be, the narrower will be the range of characteristics to which it will apply, the less we shall be able to exclude, and, consequently, the more complex the model will become and less precisely it will fit the rich diversity of objects which it is meant to summarize, and consequently the less of a model . . . it will necessarily be [p. 14].

The most rigorous and universal of all models is that of mathematics which operates at the highest level of abstraction from nature. As the sciences move down the scale, Comtean fashion, they become richer in content and less susceptible to the rigor and quantification applicable to the so-called "hard" sciences. History for Berlin is at the bottom of this scale.

Additional contrasts are made. Differences in events are the main interest of historians, whereas their regularities are the concern of scientists. Science also employs inductive and hypothetico-deductive measures, whereas historians rely on an amalgam of observation, intelligence, imagination, and empirical insight. Experimentation is also the trademark of science. The natural scientist provides scientific models of explanation; the historian relies on *Verstehen* (Nagel, 1963) to reconstruct the historical circumstance.

Having outlined his method of history, Berlin exclaims that "history is what historians do" and cautions the reader to remember the "inner" and "outer" natures respectively of historical and scientific studies. The scientist seeks factual knowledge or the "outer" conditions. The historian probes understanding or the "inner" events. Berlin thus attempts the removal of a rigorous scientific ground from history and relegates much of the past to an inquiry necessarily of a subjective nature.

An evaluation of Berlin's work in the light of educational history[2] is difficult. A number of issues, however, are deserving of comment. The depth, complexity, and richness of historical events have not been

[2] One might distinguish between history of education which emphasized the institutional study of education, and educational history which views education in its informal setting as well as in its formal institutional aspects.

overlooked and seem to reflect the crux of Berlin's methodological position. It is certainly true that the variables in any historical situation and their effect upon groups or individuals are often not readily known or even open to possible discovery. The reasons underlying the unpredictable as well as the purposeful actions of that totality, man, remain enmeshed in the "subconscious," forever hidden from precise empirical tools of investigation. The problems here are manifold. Can the historian *know all the facts?* What do we mean by "know"? What do we mean by "facts"? The value of *Verstehen,* retrodiction, and other devices thus become indispensable to the historian for completing his inquiry. Berlin is aware of the values and cultural biases of this historian, a model or models which affect the investigator's selection of facts and their interpretation. Even the "positivist" historian interprets and organizes facts on a set of criteria that may rest on "irrational" grounds.

The admixture of philosophy and methodology in Berlin's work, however, is in need of further clarification. He has told us how history is sometimes but not always written; he has further delimited himself to an acceptance of the *status quo* in historiography. Analysis of his work raises questions as to the possibility of improving historical techniques and as to the possibility of increasing the sophistication of the critical philosophy behind the historical facade. "The Concept of Scientific History" is not so much a denial of the value of scientific techniques applied to history, as the denial of the value of a scientific ideal in the study of history. There is a tendency to decry the importance of "precization" within methodological processes, the adoption of which Berlin assumes would destroy the independence of history as a separate discipline. As an alternative to the scientific ideal, Berlin offers a type of "historian's mystique" and insight which in part apparently rests on a "qualitative, quasi-intuitive form of thinking [p. 17]."

Historical happenings are a rich complexity of factors, unclassifiable and unique. The historian must employ his ingenuity, particularly his subjective insight into the affairs of men, to introduce into the thread of historical thought a semblance of continuity and probability. Scientific tools are ancillary for the historian. These views are indicative of the free-will theorist, the opposite of which is the determinist, who for Berlin invariably exhibits an empirical frame of mind.

Another of Berlin's concerns is the distinction of the natural sciences from history, insofar as the former employ a language often technical in nature and devoid of the normal moral and psychological concepts of daily speech. He then separates, for example, the history of technology from general history on the above basis. This distinction raises a twofold problem: first, whether or not there can be a history of technology or of education, empty of the cultural context, as suggested by Berlin;

and second, if so, whether or not Berlin's criterion for such a separation is sound.

A fundamental difficulty with Berlin's work, and almost any discussion of a philosophical or methodological nature is the matter of definition, the lack of which results in vagueness and deception. Berlin fails to heed the need for definition and thus introduces such crucial distinctions as "insight," "scientific reasoning," "self-consciousness," and others without delineation by either specific identification or even contextual clarification, thereby leaving what is meant to the reader's own interpretation. This leads to contradiction. Berlin, for instance, employs the term "science" both as a technical device and as a separate discipline, not bothering to elucidate what he means. He then criticizes the value of scientific explanation for historical analysis but admits its value as an ancillary procedure.

Similarities and differences between history and natural science are also key concerns. Berlin's emphasis on the differences, to the exclusion of similarities, tends to create a greater gulf than need exist between these approaches to knowledge.

> The purpose of historians, as has often been repeated, is to paint a portrait of a situation or a process, which like all portraits seek to capture the unique pattern and peculiar characteristics of its particular subject; not to be an x-ray which eliminates all but what a great many subjects have in common [p. 19].

The claim that historical events are unique and hence unclassifiable is countered by Q. Gibson (1960) who believes that the difficulty rests on:

> . . . a simple verbal confusion which can be cleared up without much difficulty. What, after all, is meant by those who insist on the uniqueness of an event? They are hardly likely to mean merely that such an event is a *particular* event, numerically distinct from others. For in this sense every event is unique, the swing of the pendulum as well as the French Revolution. And anyway, the particularity of events of whatever kind is not something which we could be said to miss when we describe them. In giving a description we do not merely name features to the particular event in question. Abstraction of features and reference to particulars may clearly go along together [p. 9].

It would seem that the extent to which we regard something as an instance of a type or as unique is a function of our point of view or interest. The need for the historian to provide reasonable explanations of events requires that he be aware of the use of both particularity and generality (Cahnman and Boskoff, 1964).

Related to this question is that of explanation. Berlin distinguishes between scientific and historical explanation on the basis that the former involves a strict set of hypothetico-deductive and inductive procedures, while the latter exhibits qualities of insight and psychological identification. This distinction seems to parallel the "outer" and "inner" meanings of history—if there are such distinctions. To suggest that in history an

explanation involving empathy is superior to one founded on principles of science is merely to emphasize a particular aspect or point of view. History, however, involves the totality of the past and requires all available techniques for its interpretation, even if such techniques are from the camp of science. Assistance can be gained by reference to W. Gallie's (1964) conception of history and explanation when he notes that explanations "in the sciences . . . play a supreme creative role: they mark the vital growing points as well as the points of positive achievement [p. 112]." He goes on to say:

> But by contrast, no one expects an historian to be an originator or unifier of the laws and theories which are exemplified in his work. What we expect from him is the ability to use other men's laws and theories, as and when they are relevant for his own purposes: which is to help us to see which is the most likely of a number of conceivable or followable developments in the difficult because imperfectly evidenced narrative that he is trying to present [p. 113]."

Gallie has moved a step further than Berlin in admitting the value of scientific explanation for history, but even here, his too rigid distinction between the function of explanation in the sciences and history tends to exclude even the possibility of employing a scientific explanation as a "story" in itself, for even a scientific explanation could be the complete story of history albeit expressed through the narrative or any other adaptable medium.

Berlin concedes that history must employ "all the knowledge and skills that we have acquired from whatever quarter," but he is quick to suggest that "a man who lacks common intelligence can be a physicist of genius." Like many in the field of philosophy, Berlin has perpetuated the unfortunate separation of man's quest for knowledge into competing, warring segments. History, for him is a separate discipline such as mathematics, biology, or sociology. What Berlin has failed to emphasize is that history employs *all approaches to knowledge* (models, if you wish) in order to obtain the best and most comprehensive view of historical events. The ensuing thoughts, I believe, are representative of the true inquirer in any endeavour:

> Human beings . . . are only on the threshold of the most preliminary steps to the mysteries of Man and the Cosmos.
>
> . . . . . .
>
> Thrown into an eternally changing universe, human beings cannot be tied by a set of rigid rules for language, thought, or action. Assuming that we, and most of our fellow beings choose to exist, and to increase our insight, perfect our knowledge about ourselves, our fate and our cosmic situation, we should never express any judgement of value or truth without carefully considering the status of present relevant research [Tennessen, 1965, pp. 40-41].

The educational historian, approaching his data with the ideal of objectivity, should have an awareness of the complexity of historical

events and the many factors affecting their presentation. In dealing with the formal as well as informal educative agencies such as the church, periodical press, and family, he should, as far as is possible, recognize his own values which could affect the interpretation of his facts. The history will be further affected by the selection of the topic of educational influence, the period of time with which the history is concerned, the availability of reliable data, and the specific purpose of this history.[3]

Providing that the educational historian does not contravene the principles of historiography (Barzun and Graff, 1957; Knopf, 1963), he should use any techniques that assist him in explaining the past. In accounting for the educational influence of the family, for example, he could refer to explanatory theories derived from sociology. By reserving the right to select any or all procedures of description and explanation, the historian is better able to interpret the complexity of educational history.

Historical knowledge thus occupies a commanding place in the researches and inquiries of mankind. Depending on one's definition, "history" is either an art or a science, or it is both, or it is neither. Most educational historians would probably agree, however, that history is the enquiry into the past of man with reference to actual educational events that have occurred. Man remains the key, for without him educational history cannot be written. The environmental and cultural forces become the indispensable setting.

Educational history, though a study in itself, frequently brings together the results of the investigations that the other studies conduct, and shapes them into a comprehensive account related to the course of historic events. Man seeks to increase his insight and perfect his knowledge of the world. This he accomplishes by using various models to describe and interpret the world. These models, depending on their goals or ends, become the interrelated disciplines of the natural, social and behavioural sciences, and the humanities—thereby assisting man in his task. The educational historian has at his disposal all of man's efforts. He should strive to improve the scientific techniques of historical research, and, being cognizant of the problems of complexity, he should make greater effort to attain historical understanding. He should remain flexible, and keeping in mind the purpose of his investigation, he should select, reject, or combine differing theories if they assist him in gaining even the most minute insight into the explanation of man's

---

[3] The educational historian's purpose is dependent on the kind of history he wishes to write. For example, he may wish to examine in depth the curriculum of a formal institution of learning in which case he might produce a monograph study. On the other hand, he may wish to explain the numerous factors affecting the development of an idea in which case he may concern himself with either a broad cultural or social history.

historical past. Thus, his awareness of what Berlin considers to be the complexity of history would be tempered with scientific ideals vital to the development of knowledge.

References

Barzun, J., & Graff, H. *The Modern Researcher.* New York: Harcourt, Brace & Co., 1957.

Berlin, I. History and theory: The concept of scientific history. *History and theory,* I, 1961. Pp. 1-31.

Cahnman, W., & Boskoff, A. (Eds.) Sociology and history: Reunion and rapproachment. *Sociology and history: Theory and research.* New York: Free Press of Glencoe, 1964. P. 4.

Gallie, W. *Philosophy and the historical understanding.* London: Chatto & Windus, 1964. Pp. 112-113.

Gibson, Q. *The logic of social enquiry.* London: Routledge & Kegan Paul, 1960. P. 9.

Gottschalk, L. *Understanding History.* New York: Alfred A. Knopf, 1963.

Nagel, E. On the method of *Verstehen* as the sole method of philosophy. *Philosophy of the social sciences: A reader.* M. Natason (Ed.) New York: Random House, 1963. Pp. 363-365.

Tennessen, H. The serio-comic encounter of clinical psychology and existential philosophy. Abbreviated version of a paper read before the International Psychology Congress at Banff, Alberta, April, 1965. Pp. 40-41.

Alberta Journal of Educational Research
Vol. XIV, No. 3, September, 1968

Optional shift performance of 40 eight-year-old suburban middle-class children was compared with that of a comparable group of North American Indian children on two types of discriminanda: concrete familiar objects and abstract pictorial material. The data showed that children from a culturally different environment, lacking in verbal experience, tend to display a retarded development of mediating responses in a concept formation task.

DONALD W. KNOWLES
The University of Victoria

and

FREDERIC J. BOERSMA
The University of Alberta

Optional Shift Performance of Culturally-Different Children To Concrete and Abstract Stimuli[1]

The native Indian population in Western Canada manifests many of the difficulties in school learning tasks exhibited by "culturally disadvantaged" students studied elsewhere. Such factors as crowded dwellings, limited education of parents, unfamiliarity with time schedules, multiplicity of dialects, and poor nutrition have characterized the environment of many Indian children. A disproportionately small number of Indian students attains more than grade eight standing compared to the non-Indian population, with less than 1% progressing as far as grade twelve (Waller, 1965).

An important factor in this attrition would appear to be the lack of development of symbolic thought behaviour required by many of the school tasks. Several recent studies have found an association between the development of abstract thinking and experiential background, generally as indicated by indices of socioeconomic status (s-e-s). John

[1] This study was supported in part by the Indian Affairs Branch, Educational Services Division, Edmonton, Alberta, and A.A.C.E.S. Grant 20066.

(1963) found that lower-class children were surpassed by higher-class children on measures of labelling and categorizing behaviour. Higgins and Silvers (1958) found that Negro children of low s-e-s background showed poor performance on both the Stanford-Binet Intelligence Scale and a purportedly non-verbal test, Raven's Coloured Progressive Matrices.

The critical effect of disadvantaged environments is often of a verbal and conceptual nature (Bernstein, 1965; Clark, 1963; Deutsch, 1963; and Taba, 1966). As an example of the importance of verbal experiences, Prehm (1966) showed that verbal pretraining leads to significantly better performance on concept acquisition tasks among a lower s-e-s group of four to seven-year-olds. One explanation of the retarded development among disadvantaged children would be that, because of lack of verbal stimulation and experience, language does not come to serve a mediating function in directing behaviour and assisting learning. This mediating function is considered to develop once facility with vocabulary and syntax have been attained (Luria, 1957). Thus, facility with the communication function of language is necessary for the development of its mediating function. In addition, Jensen (1966) has hypothesized that an "arousal threshold" of mediating responses exists which concerns the tendency for particular stimuli to elicit verbal behaviour. This threshold is seen as being "directly related to the amount of practice in, and reinforcement of, verbal learning behavior and other forms of verbal coping with the environment (Jensen, 1966; p. 104)."

An important element in the study of mediating responses would appear to be the S's familiarity with the stimuli expected to elicit the mediating responses. It would be expected that familiar stimuli would more likely have a hierarchy of mediating responses available than would unfamiliar stimuli (Osgood, 1964). Some research findings in this direction have been indicated. For example, Sigel and McBane (1966) found that lower-class children had difficulty in classifying two-dimensional representations, even full-sized photographs of objects, despite consistent success in classifying the objects themselves. The introduction of familiar objects in tasks calling for mediational responses would provide important comparisons with performance on tasks using less familiar, two-dimensional figures. It would be expected that Indian children would more readily have available mediating responses for objects which are familiar to them and with which they have had experience in labelling.

The investigation of mediating responses has recently been conducted on the basis of performance of Ss on reversal and nonreversal shift tasks (Buss, 1953; Harrow & Friedman, 1958; Kendler, *et al.*, 1959, 1962, 1966; and Sanders, Ross & Heal, 1965). Here the S is required to discriminate initially between stimuli which differ simultaneously along two

dimensions. Once this initial discrimination has been learned, the S is required to shift his manner of responding. In the reversal shift, the same dimension is maintained but the choice must be reversed. In the nonreversal shift, the S must choose a stimulus from the previously irrelevant dimension. A simple example would be a task involving cups of two sizes, large and small, simultaneously varying on the dimension of brightness, white and black. The S would be required to select the larger cup of two presented him, regardless of whether it was white or black. A reversal shift would require him to respond to the small cup, that is, to reverse his initial response. A nonreversal shift would involve drawing the positive instance from the previously irrelevant dimension (e.g., the black cup). Kendler and Kendler (1962) suggest that if mediating responses are available, the reversal shift is easier since the same mediating response is appropriate and only the overt response must be changed. By comparison, nonreversal shift learning requires that both the mediating response and the overt response be changed since the initial dimension does not maintain its relevance.

A series of experiments by Harrow and Friedman (1958), Buss (1953), and Kendler and D'Amato (1955) indicated that college students execute a reversal shift more rapidly than a nonreversal shift. Experimentation with young children (Kendler and Kendler, 1959) showed that, when considered as a group, kindergarten children learn the two types of shift at approximately the same rate. When the children were considered as two groups, fast and slow learners, on the basis of initial discrimination learning, it was found that fast learners performed as predicted by mediational theory in that they found the reversal shift easier. Since the other groups found the nonreversal shift easier, it was assumed that they were operating in a nonmediational, single unit S-R manner.

The inferred developmental effect on the acquisition of mediational responses was studied by Kendler, Kendler, and Learnard (1962). The experimental procedure was changed to that of "optional reversal shift learning" which presented a choice of responding in either a reversal or nonreversal manner after initial discrimination learning. As hypothesized, the proportion of children who responded in a reversal manner increased with age. The percentage rose gradually from 37.5% at age three to 62.5% at age ten.

In the present study, the optional shift procedure was adopted to investigate the presence of mediational responses among Indian children. It was considered preferable to the older procedure, which compared number of errors in reversal and nonreversal shift tasks, since it examines the way in which S learns, rather than postulating the possible cause of failures. In addition, the optional procedure controls for confounding effects of partial reinforcement (Harrow and Friedman, 1958).

The primary purpose of the study was to examine the relationship between cultural difference and performance on reversal-nonreversal tasks as an indication of the development of mediational responses. More particularly, it attempted to investigate the responses of two culturally-different groups on two types of tasks, the first characterized by concrete familiar objects and the second by abstract pictorial stimuli. The performance of a sample of Indian children on the two types of discriminanda was compared with that of children from suburban middle-class homes. Since preference for the reversal shift is assumed to be associated with performing in a mediated fashion, it was anticipated that fewer Indian Ss would fall in this group because of their lack of verbal facility. Furthermore, it was anticipated that more Indian Ss would be reversers when concrete stimuli, rather than abstract forms, were presented as a result of greater familiarity with the actual objects.

Method

Subjects

The Ss were 80 school children drawn from two types of experiential backgrounds. Forty children were from suburban middle-class homes. Parental occupations of this group were generally in the professional and skilled trades areas. The other 40 Ss were drawn from three Indian reservations within the Edmonton Agency of the Department of Indian Affairs. The Indian population studied was judged by Indian Affairs personnel to be representative of the majority of Alberta Indians residing on reserves in terms of degree of acculturation, s-e-s, and educational level. Few parents had permanent employment; source of income was generally welfare payments or band funds from sale of mineral rights. The Indian students came in equal proportions from Cree and Stoney language backgrounds although in most cases parents spoke English as well. Ages of the children ranged from 8 years, 0 months to 8 years, 6 months, with mean ages of 99.9 and 99.3 months for the Indian and middle-class groups, respectively.

Apparatus

The apparatus for presentation of the abstract stimuli was similar to that described by Kendler, Kendler and Learnard (1962). Pairs of squares were presented on an apparatus that permitted the child to indicate his choice by pushing a lever, receiving a candy reward if he made the correct choice. The discriminanda consisted of four squares, differing in size and brightness, mounted in pairs on a mid-gray pasteboard card. Two squares were small (1 sq. in.) and two large (3 sq. in.); two were black and two white. The squares were paired in such a way that they varied simultaneously on the two dimensions. Thus, the small black was paired with the large white and the small white with the large black.

The apparatus for presentation of concrete objects consisted of a box, the face of which was 24″ x 18″, with an opening 12″ x 7½″ which was closed between trials by the lowering of a screen. Discriminanda consisted of four tin cans, judged to be familiar to all Indian children, which differed in size and brightness. Two cans were large (4″ tall x 2½″ diameter) and two were small (2¾″ tall x 2″ diameter). One can of each size had a black label; the other two had white labels. The cans were presented in a similar pairing to the squares so that members of each pair varied simultaneously on two dimensions. The cans were presented 6″ apart. A candy was placed under the correct can between trials.

Procedure

The Ss were drawn from the total within the school in the appropriate age group on a random basis with no attempt to match groups on variables other than sex. Equal numbers of boys and girls were placed within each of the comparison groups. Preliminary analysis of the data indicated, however, that no differences existed in the performance of boys and girls. To assure that no seriously retarded children were included, Bender's Visual Motor Gestalt Test was given. All children performed above the lower level of the average range as indicated by the Koppitz scoring method (Koppitz, 1964).

Each S was seen individually. He was seated at a table and told that he was to take part in some learning games, the first of which required him to copy some figures. The Bender-Gestalt Test was then administered according to standard instructions (Bender, 1946). Upon

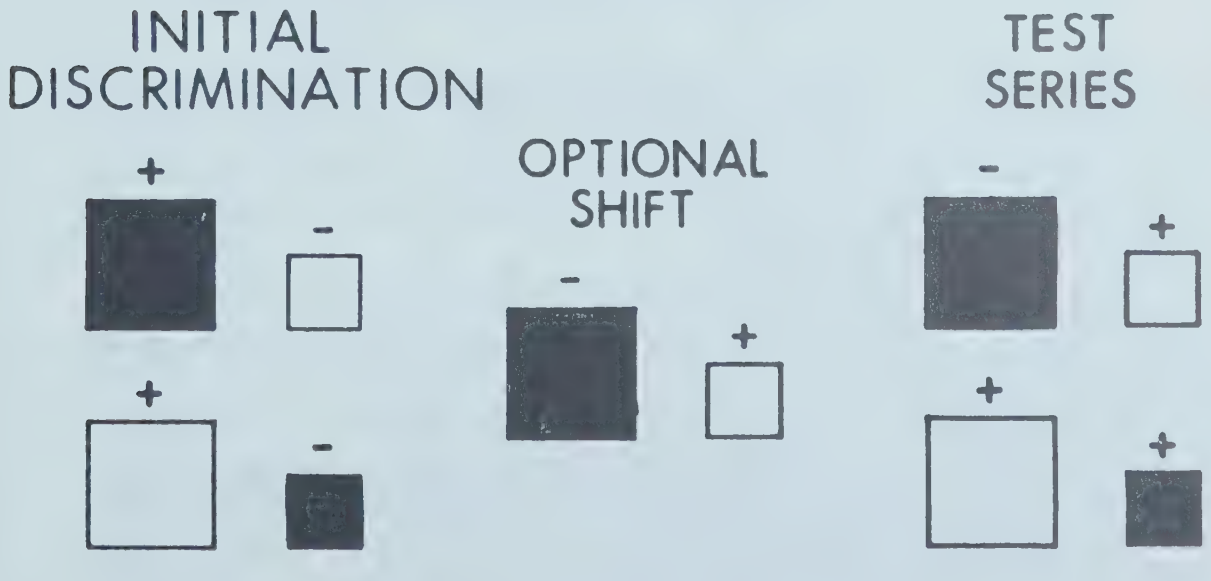


Fig. 1. Example of the abstract stimuli used and one of the reinforcement patterns.

the completion of the drawing task, S was told he was going to play another learning game and was conducted to a second table on which the mediation test apparatus was set out. Previously, Ss had been randomly assigned to discriminanda groups (concrete or abstract) and to initially relevant dimension groups (size or brightness).

The optional reversal shift procedure was similar to that described by Kendler and Kendler (1966). The test procedure consists of three stages as illustrated in Figure 1.

Stage I—Initial Discrimination: The pairs of discriminanda were presented alternately in a prearranged series so designed that: (*a*) each pair appeared equally often but no more than twice in succession, and (*b*) the correct stimulus appeared equally often on the left and right but no more than twice in succession in any one position. To discourage position habits, the stimuli were not changed until a correct response had been made. A response to one of the pairs was always correct and was rewarded with a candy. For half of the Ss, size was the relevant dimension; for the other half, brightness. Criterion for initial discrimination learning was 9 correct out of 10 consecutive responses, immediately following which the next stage was presented.

Stage II—Optional Shift: Only one of the stimulus pairs was presented and the previous reinforcement pattern was reversed. Both dimensions, size and brightness, were now relevant. Thus, in the illustrated example, S could learn the shift by responding to "small" or to "white" or to both indiscriminately. Criterion was again 9 correct out of 10 consecutive responses.

Stage III—Test Series: Immediately following the optional shift learning, both pairs were presented in the same prearranged order used in the initial stage. The pair not presented during the optional shift was considered as the "test pair." Responses to this pair determined the basis of response during the optional shift. Responses to *either* stimulus in the test pair were rewarded. The pattern of reward for the other pair followed that used during the optional shift to maintain the shifted response pattern. Each pair was presented 10 times in the alternating sequence used initially.

Following the 20 presentations in the test series, S was shown the "optional shift" pair and asked to explain which one was the "winner" in an attempt to learn more of the verbal process involved. At the end of each session, S was asked to keep information about the games secret until others had had their turns.

The accuracy of each response during the optional shift procedure was recorded. From the score sheet, number of trials during initial discrimination and optional shift trials were compiled for each S. On the basis of responses to the *test pair,* S was classified as a reverser if he made 8 or more responses in a manner opposite to the way he had responded during initial training (Kendler and Kendler, 1966). Nonreversers included all Ss who made 8 or more responses in the same manner as they had during initial training. All other Ss were considered as belonging to the inconsistent group.

The explanation provided by S regarding the "winner" was recorded verbatim for later classification into one of the following categories:

(a) verbalized correct dimension—responses to one or both of the dimensions which apparently guided S's responses in the optional shift and test series; (b) verbalized incorrect dimension—responses that were inconsistent with overt choice; (c) irrelevant—verbal explanations which did not mention either dimension on which the stimuli differed; (d) no verbalization.

The percentage of Ss within each of these four classifications was considered in terms of choice on the test series—reversal, nonreversal, and inconsistent.

Results and Discussion

Number of Reversers

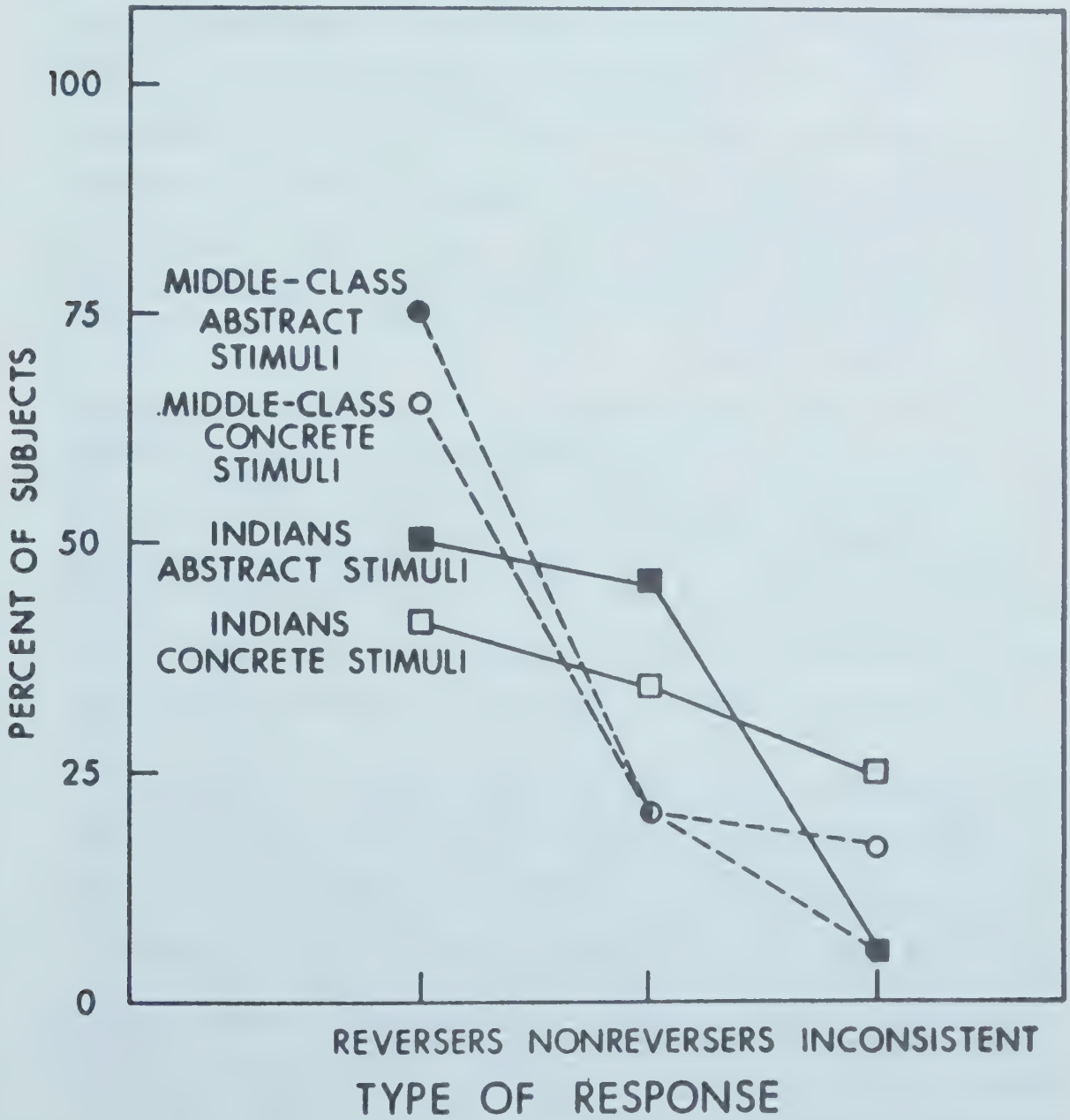


Fig. 2. Percent of subjects from each cultural group at each level of responding on test series.

As indicated in Figure 2, there was a consistent tendency for a lesser proportion of Indian Ss to respond in a reversal manner. Contrary to expectations, use of concrete stimuli did not lead to a greater number of reversers in the Indian group. Analysis of the relative influences of cultural group, type of discriminanda, and initially relevant dimension by multpile classification χ^2 (Sutcliffe, 1957) indicated that significantly more middle-class children (n=28) reversed as compared to Indian children (n=18). The obtained chi-square value for the cultural factor was 5.1 (df=1, $p<.025$). None of the associations with other factors nor interactions among factors was statistically significant.

The lack of significant association between type of discriminanda and number of reversers appears to be related to the disproportionate number of Ss responding inconsistently in tasks using concrete stimuli (Figure 2). It would serve to suggest that a certain amount of generality exists for the labelling of learners as reversers or nonreversers on the Kendler task. While it has been assumed that different types of discriminanda produce similar results in variations of the mediation test, for example Wolff (1967), no empirical support has been offered. In the present study, modifications in the procedure did not produce significant changes in the number of mediators in a group, at least as long as the stimulus values remained the same.

The lack of effect from changing the discriminanda is nevertheless surprising. It was assumed that Ss have had more experience in labelling and dealing with cans than squares. One possible explanation of the data would be that, because the cans were actually handled, size would become a predominant dimension and would interfere with learning the dimension of brightness and maintaining it during the shift learning. The results indicate, however, that the pattern for number of reversers on brightness and size dimensions is similar for the cans and the squares. A second explanation would return to Osgood's (1964) notion of a hierarchy of mediational responses being associated with a particular stimulus. It could be that, while mediational responses for brightness and size dimensions are associated with cans, other responses are more strongly associated and their elicitation interfered with performance in the present task. By contrast, mediational responses for brightness and size would be predominant for the squares. In partial support of this interference explanation is the disproportionately large number of Ss in both groups who responded in an inconsistent manner on the test series with the cans (Figure 2). An improvement in the procedure would be effected by selecting materials which differentiate between Indian and middle-class groups on the extent to which they elicit appropriate verbal responses.

Verbalizations

The Ss explanations of choice of winning can or square are reported in Table 1. An interesting result, in view of the commonly-held impression of Indian children as being reticent with adults, was that Indian Ss did not show withdrawal from verbal expression in this task. It was necessary, however, to add the question, "Which is the *good* one," when little response was made to the initial question about "winners." This extended questioning was required for about one-half the Indian Ss. Middle-class Ss showed a marked tendency to elaborate spontaneously upon their answer, e.g., ". . . the white one. At first, the black one was right but then it switched."

TABLE 1

PERCENTAGE OF SUBJECT'S GIVING TYPES OF VERBAL EXPLANATIONS AS A FUNCTION OF THEIR CHOICE ON THE TEST SERIES

			Verbalization			
Choice	Group	n	Correct	Incorrect	None	Irrelevant
Reversal	Middle-class	28	85.7	14.3	----	----
	Indian	18	77.7	16.7	----	5.6
Nonreversal	Middle-class	8	75.0	----	25.0	----
	Indian	16	63.0	37.0	----	----
Total	Middle-class	36	83.3	11.1	5.6	----
	Indian	34	70.6	26.5	----	2.9

Note--Inconsistent responses were given by 4 middle-class and 6 Indian children.

The majority of reversers in both the middle-class and Indian groups made correct verbalizations. Total number giving correct verbalizations was greater for middle-class children than Indian, but the difference was not statistically significant (χ^2=3.05, df=3, N.S.). A lesser number of nonreversers correctly explained their choice behaviour in both the cultural groups, with a somewhat greater proportion of middle-class than Indian Ss falling in this category. Interestingly, the nonreversers in the middle-class group who could not give correct explanations all fell within the "none" category. The similar group of Indian Ss all gave incorrect explanations. All Ss in both cultural groups who were in the in-

consistent choice group were able to give explanations. Since no dimension was consistently employed in their choice behaviour, it was impossible to classify these verbalizations as being correct or incorrect.

Comparisons of verbal explanations of choice behaviour would appear to provide important information about the relationships among overt language, test performance, and cultural group. It is significant that, despite general reticence to verbalize, Indian Ss did provide verbal explanations of their choice in this procedure (Table 1). Middle-class Ss tended to explain their behaviour more accurately than Indian Ss. An important finding to consider is the verbalizations of Indian Ss who were classified as nonreversers. When they were asked to explain which was the winner following the task, this group behaved in an almost random manner. The middle-class group of nonreversers did not show a similar tendency. The discrepancy would suggest that Indian Ss, particularly the nonmediators, do not have the same tendency to use verbal labels with non-verbal stimuli as compared to urban middle-class Ss in agreement with the higher arousal threshold hypothesized by Jensen (1966).

Number of Trials to Criterion

The relationship between reversal preference and speed of learning suggested by Kendler *et al.* (1962) leads to the expectation that fewer trials to criterion would be found for the reversal group. Similarly, as-

TABLE 2

NUMBER OF TRIALS TO CRITERION FOR INITIAL AND OPTIONAL SHIFT STAGES

		Initial Discrimination			Optional Shift		
Choice	n	Mean	Median	Range	Mean	Median	Range
Indian							
Reversal	18	23.7	15.5	9-102	13.1	11.0	10-24
Nonreversal	16	21.1	17.5	10-56	14.4	14.0	11-20
Inconsistent	6	15.2	15.0	10-21	13.5	13.5	11-16
Middle-class							
Reversal	28	23.3	14.5	9-119	15.2	12.0	10-35
Nonreversal	8	23.4	16.0	10-51	13.3	12.5	10-20
Inconsistent	4	18.2	12.0	10-39	15.7	11.5	10-30

suming that children from advantaged cultural backgrounds are more likely to have relevant mediating responses available, one would expect that fewer trials to criterion would be needed by the middle-class group.

The number of trials required by reversers to reach criterion was found to be essentially the same as the number required by Ss in nonreversal and inconsistent groups (Table 2). And there was no association between speed of learning and cultural group. These results are dissonant with those reported by Kendler, *et al.*, but closer examination of their results reveals that in the earlier study totals across five age levels, 3½ to 10½ years, were compared. Since the greater proportion of the nonreversal group would have been drawn from the younger age groups, the effects of nonmediating and age on the speed of learning were confounded. From the findings of the present study, it would appear that the error score was related more strongly to the age factor.

The lack of relationship between speed of learning and type of choice behaviour is not particularly surprising in view of the specific nature of the task and the small number of average errors to criterion. Mediating responses which facilitate learning are more important in complex tasks, and in tasks not as similar to original learning as the optional shift phase of the present study. The problem of finding appropriate tasks to test the hypothesized advantage of being a mediator appears to be an important one. A recent study (Wolff, 1967) did compare performance of mediators and nonmediators, as determined by the Kendler procedure, on another task supposedly calling for mediation, but few guidelines exist for choosing a concurrent task that would be appropriate. Kendler *et al.* (1962) found an association between speed of learning and reversal choice; however, their design confounded age and choice behavior. Buss (1953) did not find any relationship in his study which used one age group of Ss, college-level students. In the present study, reversal Ss learned faster on the optional shift than on initial discrimination, but so did all other Ss. This comparison, although reported by Buss and Kendler *et al.*, is difficult to interpret since the optional shift stage employs only one pair in contrast to the two pairs used in the initial discrimination.

To sum, it appears that Indian children of this age do not reverse to the same extent as middle-class children. Substitution of the supposedly more familiar and concrete stimulus material did not lead to increased reversal choice for this group.

Reversal response has been used as an index of verbal mediation. An additional contributing factor to the Indians' lack of reversing could be inappropriate attentional responses (Fellows, 1968; Kendler, Glucksberg, and Kenton, 1961; McConnell, 1964) to the important dimensional cues of the problem.

The results indicate that the general approach employed can lead to valuable observations about a group of culturally-different children.

The shift procedure did not meet with resistance on the part of any S and has the advantage of being couched in a meaningful framework so that results suggest important descriptions of intellectual development. The procedure's lack of emphasis on verbal expressive ability was a particular advantage for the study of Indian children. Additional investigation of this deficit in covert language among Indian children, examining the differential behaviour of mediators and nonmediators on other tasks, is needed.

References

Bender, L. B. *Instructions for the use of Visual Motor Gestalt Test.* American Orthopsychiatric Association, 1946.

Bernstein, B. A socio-linguistic approach to social learning. In J. Gould, (Ed.), *Penguin survey of social sciences,* 1965. Middlesex: Penguin Books, 1965, 144-168.

Buss, A. H. Rigidity as a function of reversal and nonreversal shifts in the learning of successive discriminations. *Journal of Experimental Psychology,* 1953, 45, 75-81.

Clark, K. B. Educational stimulation of racially disadvantaged children. In A. H. Passow, (Ed.), *Education in depressed areas.* New York: Columbia University, 1963, 142-162.

Deutsch, M. The disadvantaged child and the learning process. In A. H. Passow, (Ed.), *Education in depressed areas.* New York: Columbia University, 1963, 163-179.

Fellows, B. J. *Discrimination process and development.* New York: Pergamon Press, 1968.

Harrow, M., & Friedman, G. B. Comparing reversal and nonreversal shifts in concept formation with partial reinforcement controlled. *Journal of Experimental Psychology,* 1958, 55, 592-597.

Higgins, C., & Silvers, C. M. A comparison of S-B and Colored Raven Progressive Matrices I Q for children with low s-e-s. *Journal of Consulting Psychology,* 1958, 22, 465-468.

Jensen, A. R. Verbal mediation and educational potential. *Psychology in the Schools,* 1966, 3, 99-109.

John, V. P. The intellectual development of slum children: Some preliminary findings. *American Journal of Orthopsychiatry,* 1963, 33, 813-822.

Kendler, H. H., & D'Amato, M. F. A comparison of reversal shifts and nonreversal shifts in human concept learning. *Journal of Experimental Psychology,* 1955, 49, 165-174.

Kendler, H. H., Glucksberg, S., & Jaston, R. Perception and mediation in concept learning. *Journal of experimental psychology,* 1961, 61, 186-191.

Kendler, T. S., & Kendler, H. H. Reversal and nonreversal shifts in kindergarten children. *Journal of Experimental Psychology,* 1959, 58, 56-60.

Kendler, T. S., & Kendler, H. H. Optional shifts of children as a function of number of training trials on the initial discrimination. *Journal of Experimental Child Psychology,* 1966, 3, 216-224.

Kendler, T. S., Kendler, H. H., & Learnard, B. Medicated responses to size and brightness as a function of age. *American Journal of Psychology,* 1962, 75, 571-586.

Koppitz, E. M. *The Bender Gestalt Test for young children.* New York: Grune & Stratton, 1964.

Luria, A. R. The role of language in the formation of temporary connections. In B. Simon, (Ed.), *Psychology in the Soviet Union.* Stanford: Stanford University Press, 1957, 115-129.

McConnell, O. L. Perceptual versus verbal mediation in the concept learning of children. *Child Development,* 1964, 35, 1373-1383.

Osgood, C. E. A behavioristic analysis of perception and language as cognitive phenomena. In R. C. Harper *et al., The cognitive processes: Readings.* Englewood Cliffs: Prentice-Hall, 1964, 184-210.

Prehm, H. J. Concept learning in culturally disadvantaged children as a function of verbal pretraining. *Exceptional Children,* 1966, 32, 599-604.

Sanders, B., Ross, R. E., & Heal, L. W. Reversal and nonreversal shift learning in normal children and retardates of comparable mental age. *Journal of Experimental Psychology,* 1965, 69, 84-88.

Sigel, I. E., & McBane, B. Cognitive competence and level of symbolization among five-year old children. Paper presented at A. P. A. Convention, 1966.

Sutcliffe, J. P. A general method of analysis of frequency data for multiple classification designs. *Psychological Bulletin,* 1957, 54, 134-137.

Taba, H. Cultural deprivation as a factor in school learning. In J. C. Gowan, & G. O. Demos, (Eds.), *The disadvantaged and potential dropout.* Springfield: C. C. Thomas, 1966, 212-231.

Waller, L. P. G. (Ed.), *The education of Indian children in Canada.* Toronto: Ryerson Press, 1965.

Wolff, J. L. Effects of verbalization and pretaining in concept attainment by children in two mediation categories. *Journal of Educational Psychology,* 1967, 58, 19-26.

Alberta Journal of Educational Research
Vol. XIV, No. 3, September, 1968

This comparative study used as its subjects two matched groups of Grades XI Social Studies 20 students. The experimental subjects had repeated the Social Studies 10 course, the control subjects had not. Although both groups obtained the same mean achievement score in Social Studies 10 the control subjects earned a significantly higher score than did the experimental group in Social Studies 20.

ARTHUR G. STOREY and B. PATTERSON
The University of Calgary

What Profit Failure

The problem investigated by the present study is but another aspect of the enquiry into the merits of failing grades undertaken by Storey and Desson (1966) wherein it was suggested that when a teacher is asked why a given student failed a certain course the responses were to the efefct that: (*a*) he did not earn enough marks to pass, (*b*) he was not ready for more advanced work, and/or (*c*) he did not put forth enough effort to warrant a passing grade.

The Storey-Desson Study investigated the third of these contentions in that it tested the motivational value of low grades to discover that artificially bonused marks had greater motivational effect. It would seem to follow, then, that if a student is failing through lack of effort an inefficient method of increasing this effort is the awarding of low grades.

The issue as to whether or not a given student earned enough marks to warrant a passing grade is a nebulous one owing to the nature and calibre of the tests used to assess achievement. Most of the achievement testing done, at least in the Alberta schools, is done with teacher-made tests and most teachers have had little or no training in test construction. As a result, the tests used are, for the most part, unsatisfactorily low in validity, reliability, and objectivity. Besides this, most of the instruments used to assess achievement in areas like social studies, the subject of concern here, have traditionally been of an essay type and hence notorious for their inadequacies both in general (Wood, 1961)

and on the Alberta scene (Storey, 1968). The reliability coefficient of teacher made essay type examinations usually approaches .50 while the standard deviation of the scores approximates 15 (Storey, 1968). Thus, the standard error of measurement ($\sigma e = s\sqrt{1-n}$) is about 10 rendering any achievement grade based on the test relatively meaningless. For example, if the earned score is 65, the true score lies somewhere between 35 and 95 if one wants to be 99% confident. That is, the student might be failing the course miserably or might be passing it with extremely high marks. The test just was not adequate to determine which. It follows, then, that any pass/fail decision based on the scores obtained through such tests is ludicrous.

No doubt it is partly the results of these inadequacies in teacher-made tests that lead Alberta's Department of Education to continue to exercise responsibility for achievement testing at the crucial grade IX and XII levels. In any case, when the average classroom teacher suggests that a student failed a given course because he did not earn enough marks to pass it he is on very unfirm ground.

The claim that students were failed in a certain course because they were not yet ready for more advanced work seems more worthy of serious investigation. However, Cronbach (1963), in commenting on the findings of a number of studies on the results of repeating grades or courses, concludes that, although repeating work theoretically provides opportunity for mastery, in practice the repeater makes less progress than the student who does not repeat. The major purpose of the present study, then, is to test the supposition that students who repeat Social Studies 10 will achieve better in Social Studies 20 than do their peers who have not repeated Social Studies 10.

Subjects

In order to obtain a sufficiently large number of Grade XI students who had repeated Social Studies 10 together with matches for them, the three lowest achieving classes in each of three randomly selected Calgary high schools were used as subjects. In December each of the resultant 242 students was given a fifty item multiple choice test covering the just completed units, Units I and II, of the Social Studies 20 course. This test, constructed by four Social Studies 20 teachers with training and experience in test construction, was item-analyzed and re-scored on the basis of its best 40 items. The minimum criteria for item retention were a discrimination index at or above .35, a difficulty index between .30 and .80, and all item distractors working. The split half reliability coefficient for the forty item test, when corrected for length was .91, the S. D. 10.79 making the standard error 3.24.

Following the administration and scoring of this test all of the subjects in the sample who had repeated Social Studies 10 were identified and

matched as nearly as possible for age, Otis I. Q., sex, and Social Studies 10 final marks, with students who had not repeated Social Studies 10. The subjects who had repeated Social Studies 10 are hereafter referred to as the experimental group while their matched peers who had not repeated it are called the control group. Table 1 summarizes the results of this matching procedure. For convenience the results obtained from the Social Studies 20 test are also included in this table. All achievement marks are given in percentages. It is to be noted that the mean age of the experimental subjects is approximately one year greater than that of the control subjects.

Treatment of Data and Conclusion

Means and standard deviations were calculated for the total group and for males and females. These data are reported in Table 1. The differences of the means were analyzed using t tests of significance and the results are reported in Table 2.

TABLE 1

COMPOSITION AND ACHIEVEMENT OF CONTROL AND EXPERIMENTAL GROUPS

Control					Experimental				
Girls									
Subject	Age	I. Q. (Otis)	S.S. 10	S.S. 20	Subject	Age	I. Q. (Otis)	S.S. 10	S.S. 20
1	16-10	104	40	55	1	18-4	103	50	65
2	16-0	107	50	65	2	16-3	104	40	43
3	15-11	105	40	58	3	16-10	106	40	43
4	16-1	113	50	66	4	16-11	110	40	38
5	16-3	101	40	55	5	16-8	103	40	38
6	16-4	99	50	60	6	17-5	92	45	18
7	16-4	107	45	53	7	17-4	111	40	50
8	17-4	96	40	50	8	18-5	106	55	35
9	16-2	112	50	40	9	16-3	104	40	48
10	16-3	105	40	40	10	16-10	106	40	33
11	16-8	120	50	80	11	17-9	111	55	45
12	16-1	105	45	40	12	17-3	116	50	40
13	16-2	107	45	40	13	17-2	108	40	63
14	16-8	100	40	43	14	17-11	111	50	48
15	16-8	119	45	50	15	17-9	119	55	65
$\bar{X}$	16-5	106.67	46.67	54.33	$\bar{X}$	17-3	107.33	45.67	44.80
SD		5.48	4.25	11.84	SD		6.00	6.55	12.25

TABLE 1 (cont'd)

Boys

1	16-8	90	40	40	1	18-4	100	40	48
2	16-7	96	50	55	2	17-9	93	45	50
3	16-6	105	45	60	3	18-3	102	50	60
4	16-5	101	50	55	4	18-0	109	55	50
5	16-10	108	40	55	5	18-0	112	45	58
6	16-8	109	50	60	6	18-7	103	45	20
7	16-7	123	40	60	7	19-0	115	50	45
8	16-2	125	50	50	8	17-1	127	45	60
9	16-1	116	45	65	9	17-1	112	45	55
10	16-5	112	45	60	10	17-10	112	55	50
11	16-5	100	45	60	11	16-10	103	45	38
12	16-2	108	45	60	12	17-2	105	55	58
13	16-7	104	40	55	13	17-11	107	45	60
14	16-1	108	45	58	17	17-0	106	45	48
15	16-4	109	40	60	15	17-0	107	40	55
16	16-7	108	50	85	16	18-0	98	50	45
17	16-4	119	45	58	17	17-7	114	40	58
18	16-10	108	45	55	18	17-3	103	45	28
19	16-9	112	40	45	19	18-1	115	50	43
20	16-11	103	45	48	20	17-9	96	45	58
$\bar{X}$	16-8	108.20	46.75	58.75	$\bar{X}$	17-9	106.95	48.75	49.35
SD		7.75	3.80	9.22	SD		7.68	4.67	10.55
Total									
$\bar{X}$	16-7	107.54	46.71	56.85	$\bar{X}$	17-6	107.11	47.00	47.40
SD		7.31	4.00	10.53	SD		7.12	5.61	11.55

An examination of the data provided in Table 2 indicates that:

1. the total experimental and control groups and the sub-groups were not significantly different in ability as measured by the Otis intelligence test nor in Social Studies 10 achievement as measured by the final grades awarded by the various Calgary high schools.
2. both of the sub-control groups and the control group as a whole scored significantly higher on the Social Studies 20 test than did their experimental counterparts.

Within the limits of the current study, then, it becomes necessary to reject the hypothesis that repeating Grade X Social Studies will prepare students better for Social Studies 20.

TABLE 2

RESULTS OF T TESTS OF DIFFERENCES OF MEAN SCORES OBTAINED BY TOTAL EXPERIMENTAL AND CONTROL GROUPS AND BY MALE AND FEMALE SUB-GROUPS

	I. Q.		S.S. 10		S.S. 20	
Group	t	p[a]	t	p	t	p
Male	0.072	NS	1.900	NS	4.240	**
Female	0.410	NS	0.690	NS	3.060	**
Total	0.250	NS	0.250	NS	5.060	**

[a]NS= non-significant

* = p<.05

**= p<.01

Discussion

It is possible to debate the conclusion reached through the current study by arguing that the two groups studied were not in fact equal. Such a contention seems weak however in the face of the general rigour and control exercised through the design of the study together with the equality in Otis I. Q. scores. It is also weak when one considers that the experimental group's mean age was approximately one year greater than that of the controls and that due to school attendance regulations, this year must have been spent in school. Also, this differential age factor coupled with equivalent intelligence scores, would suggest greater maturity and experience on the part of the experimental subjects. It would seem, then, that, since the experimental group was older, had been longer in school, was equal in intelligence to the control group, and had had one school year longer in social studies, any advantages inherent in repetition had been more than counteracted through loss in motivational factors.

What profit failure, then? On the credit side, repetition may have been responsible for getting these subjects through Social Studies 10, but whether or not repetition increased their knowledge of this course material is in doubt in that any increase in marks obtained in the repetition year may have been due to the inadequacies of the tests used to assess it. Or, it may have been due to a "policy" of not failing a student twice in the same course.

The least that can be said on the debit side of failure is that, insofar as the present study could determine, it did not prepare these subjects better for new work. Indeed they achieved more poorly in Social Studies 20 than did their peers who were not required to repeat Social Studies 10.

At worst many questions could be asked concerning the cost of failure and repetition. Any such list of questions might query the cost:

a. in teacher and student time,
b. in school facilities,
c. in student and perhaps teacher motivation,
d. in student dropouts (only 35/242 low achievers were repeaters suggesting that a disproportionate number of failures had quit school),
e. in student self-development and attitudes,
f. the cost due to possible error and injustice resulting from test inadequacies, and
g. the cost due to possible scapegoating (Repeaters averaged one year older than their non-repeater peers suggesting a history of failure. There were also 20 male failures as compared to 15 female in the present sample where the total number of females [133] was greater than the males [109]).

In any case, if repetition does not accomplish whatever it is that it set out to accomplish, and the present study is one of many supporting the hypothesis that it does not (Cronbach, 1963), perhaps a careful second thought should be given before any individual is required to repeat any school grade or course.

References

Cronbach, L. J. *Educational psychology.* (2nd ed.) New York: Harcourt, Brace, & World, 1963. P. 474.

Storey, A. G. Some evidence in "The Essay Item Case." *Journal of Education Research,* 1968 (April), Vol. 61, No. 8, 351-354.

Storey, A. G., & Desson, G. Achievement as a function of assigned grades. *The Alberta Journal of Educational Research,* 1966 (Dec.), Vol. XII, No. 4.

Wood, D. A. *Test construction.* Columbia, Ohio: Charles E. Merrill, 1961. Pp. 99-102.

Alberta Journal of Educational Research
Vol. XIV, No. 3, September, 1968

A selection of achievement, intellective, and personality tests was administered to freshmen university students for the purpose of identifying variables related to first year success. Correlation and regression analyses were conducted on the resultant data. The results show high school average to be the best predictor while the other variables are shown to possess little predictive utility.

R. C. CONKLIN and D. G. OGSTON
The University of Calgary

Prediction of Academic Success for Freshman at the University of Calgary[1]

During registration week of the 1966-67 academic year, freshmen students at the University of Calgary were required to complete a battery of various achievement, intellectual, and personality scales. The students appeared for testing according to a schedule based upon the alphabetical order of their surnames. Consequently, six separate groups of students were tested at different sessions throughout the week. Each group was given a test battery of a somewhat different composition. The purpose of the testing program was to identify, where possible, any variables which might be related to freshman success. Because achievement and intellective tests have been shown (for example Brown and Dubois, 1964; Payne *et al.*, 1966) to account for only a small portion of the variance, the inclusion of several non-intellective tests was deemed desirable. In addition, some of the scales had never been previously used in a validity study of this kind.

Psychometric Instruments

The instruments used included both the well-known standardized tests and relatively untried research scales. A brief description of each will be given in this section.

1 A complete copy of this research is available from the authors.

Cooperative Academic Ability Test (CAAT)

This test is published by the Educational Testing Service (ETS) and is designed to obtain estimates of verbal and mathematical ability for advanced secondary school students. Each part contains fifty items with multiple choice answers for each. Time allowed is 20 minutes for each part. Correlations between CAAT scores and school achievement as reported in the manual tend to be about .50. In addition, correlations with intelligence tests are moderately high.

Cooperative English Test (CET)

This test, also designed and published by ETS, measures achievement in two general areas: written expression and reading. In effect, there are two tests which may be given independently. English expression breaks down into two scales, effectiveness and mechanics; reading comprehension also has two scales, vocabulary and reading. The latter reading comprehension scale has two subscales: level of comprehension and speed of comprehension. In this study three of the scales were used: (*a*) vocabulary, (*b*) level of comprehension, and (*c*) speed of comprehension. Many predictive validity studies are reported in the manual and the correlation of CET scores with various criteria of academic success yields coefficients ranging from .22 to .60, the median correlation being about .40.

The California Study Methods Survey (CSMS)

This is a self-report inventory, published by the California Test Bureau, designed to reveal the nature of the study methods and attitudes of the student. It consists of 150 standardized items which purport to reflect differences in study methods and attitudes between high achieving and low achieving students. The inventory yields four scores: (*a*) attitudes towards school (attitude), (*b*) mechanics of study 1 and 2, (*c*) planning and systems (p & s), and (*d*) verification. Only the first three scores were used in this study. The student responds "yes" or "no" to the items. Data reported in the manual indicate correlations between .32 to .57 between these scales and scholastic achievement.

Eysenck Personality Inventory (EPI)

This test (Eysenck and Eysenck, 1963) is designed to measure the personality factors introversion-extraversion (I-E), and neuroticism (N). The inventory yields a lie (L) scale which is response validity factor. There are a total of 57 items: 24 I-E items, 24 N items, and 9 L items. They are responded to by the student with "yes" or "no." No data are reported in the manual which would provide evidence of the relationship between the EPI scales and academic achievement.

Costello-Comrey Need Achievement Scales

These scales have not as yet been published and are still in an experimental form. Two scores are available which may be described as achievement via effort and achievement via the emulation of success. It has been suggested that the first scale is related to a need to do well, while the second implies a need to be successful with the added feature of fear of failure. High scores indicate a need to achieve. The effort scale is composed of 10 items while the emulation scale has 14 items. They are posed as questions and answered "yes" or "no" by the examinee.

Taylor Manifest Anxiety Scale

Though widely used, this scale (Taylor, 1953) has not been published. The scale was devised to test the hypothesis that increased drive tends to lower the efficiency of learning at all difficulty levels. The assumption is that high drive states manifest themselves as anxiety. Fifty items from the Minnesota Multiphasic Personality Inventory which purport to measure anxiety were selected. Items are answered "true" or "false" and a high score indicates high anxiety. Research has consistently shown that measures of anxiety correlate negatively with academic achievement. One such study (Kelly, Hunka, and Conklin, 1965) showed that two measures of anxiety correlated -.26 and -.11 with university average.

Internal-External Control of Reinforcement

This scale, designed by Rotter (1966), measures general expectancies in learning situations as to whether reward or success in situations is dependent upon the individual's behaviour or is controlled by external forces. The scale gives one score which is a measure of external choices. Each of the 23 items is a forced choice of two alternatives. Both theory and research with this scale would indicate that correlation with academic achievement should be negative and significant (Lefcourt, 1966).

Pittsburgh Scales of Social Extraversion-Introversion and Emotionality

This test (Bendig, 1962) was developed to provide more reliable measures of the second order factors of extraversion and emotionality. The two scales report measures of introversion-extraversion (I-E) within a social context and general emotionality (E). For most purposes the scales are equivalent to those of the Eysenck Personality Inventory. The emotionality scale consists of 30 items and the introversion-extraversion scale of 30 items. The examinee answers "true" or "false" to each item.

Procedure

As well as using measures obtained by administering the above tests, each student's high school average on the grade 12 departmental exa-

minations was obtained. Students were then followed up at the end of their first year at university and the relationships between all variables and the first year university average was studied. Complete data were available for 639 students. Initially, means, standard deviations, and intercorrelations were obtained for all variables for each of the six samples. Linear step-wise regression analyses were also computed. A step-wise method of analysis has the feature of selecting the best predictor, searching for the second best predictor, and so on. Beta weights, partial correlations, and multiple correlations with the criterion are available. In this study all variables, including the high school average, were classed as predictors while the criterion was first year university average.

TABLE 1

CORRELATIONS OF PREDICTOR VARIABLES WITH FIRST YEAR UNIVERSITY AVERAGE

PREDICTOR VARIABLES	1	2	3	4	5	6
High School Average	.34	.42	.49	.55	.52	.60
Bendig: Intro-Extra	.00	-.30	--	.16	-.11	--
: Emotionality	.02	.22	--	.18	.02	---
CAAT: Verbal	--	--	.11	.21	--	.22
: Math	--	--	.14	.08	--	.23
: Total	--	--	.14	.18	--	.29
CET: Vocabulary	.23	--	--	--	.45	--
: Level of Comprehension	.32	--	--	--	.57	--
: Speed of Comprehension	.36	--	--	--	.51	--
Costello-Comrey: Effort	.20	.18	.12	-.02	-.07	.02
: Emulation	-.13	.07	.22	-.12	-.04	.00
CSMS: Attitude	--	.06	--	.25	--	.31
: Mechanics 1	--	-.19	--	--	--	.18
: Mechanics 2	--	.04	--	.22	--	.07
: Planning & Systems	--	.01	--	.13	--	.22
: Total	--	.00	--	.18	--	.25
Eysenck: Intro-Extra	-.10	-.22	-.17	-.17	-.10	-.15
: Lie	-.11	.17	.00	-.01	-.02	.10
: Neuroticism	.05	.25	-.09	.17	.09	.09
Rotter: Internal-External	.04	-.01	-.07	-.09	.08	.00
Taylor: Manifest Anxiety	-.04	.23	--	.13	-.02	--
SAMPLE SIZE	129	95	135	106	105	69

Results

The correlation of each variable with first year university average, for each of the six samples, is reported in Table 1.

The results of the regression analyses reported in Table 2 include the predictor variable, its raw score beta weight, the constant of the equation, and the multiple correlation with the criterion.

TABLE 2

CONSTANTS, RAW SCORE BETA WEIGHTS AND MULTIPLE CORRELATION COEFFICIENTS OF VARIABLES PREDICTING FRESHMAN SUCCESS

PREDICTOR VARIABLES	FRESHMAN SAMPLES					
	1	2	3	4	5	6
High School Average	.201	.289	1.012	.751	.716	.890
Bendig: Intro-Extra	.211	-.284	--	--	-.121	--
: Emotionality	-.143	.291	--	--	--	--
CAAT: Verbal	--	--	-.194	.104	--	.297
: Math	--	--	--	-.102	--	--
: Total	--	--	--	--	--	-.168
CET: Vocabulary	.079	--	--	--	.533	--
: Level of Comprehension	.194	--	--	--	.082	--
: Speed of Comprehension	.161	--	--	--	--	--
Costello-Comrey: Effort	1.355	1.005	.704	.310	1.075	-.383
: Emulation	-.214	-.395	.998	-.892	--	--
CSMS: Attitude	--	-.360	--	.098	--	.363
: Mechanics 1	--	-1.917	--	--	--	-.285
: Mechanics 2	--	-.569	--	--	--	--
: Planning & Systems	--	--	--	-.464	--	.208
: Total	--	.630	--	.221	--	--
Eysenck: Intro-Extra	-.318	-.158	-.487	--	--	--
: Lie	-.758	--	-.552	-.779	-.377	.714
: Neuroticism	.309	.274	-.259	.554	--	--
Rotter: Internal-External	.232	.280	--	-.380	.430	.241
Taylor: Manifest Anxiety	-.056	.218	--	--	--	--
CONSTANT	25.092	38.024	-5.881	5.523	-6.652	-15.122
MULTIPLE R	.526	.677	.581	.635	.599	.685

The results of intercorrelation analysis indicate that the correlation of high school average and freshman success is consistently higher than the correlation of any other variable with the criterion. Correlation co-

efficients range from .34 to .60, the median being .49. The only other test which, when correlated with the criterion, produced correlations approaching those for high school average and academic success was the CET. These coefficients ranged from .23 to .57, the median being .41. In regression analysis it was found that although the CET correlated higher than all variables with success (except for high school average) it added nothing to the predictive efficiency because the CET did not offer any unique contribution after the contribution of high school average. Two other standardized tests, commonly used for predictive purposes, had very low correlations with the criterion. The CSMS scores and freshmen success scores correlated from .00 to .31, the median r being .15. The CAAT had measured correlations of .08 to .29, with a median of .18. The magnitude of these median correlation coefficients reflects little predictive utility of the instruments.

Regression analyses indicate that a student's high school average is the single best indicator in the present study. In five of the six samples its correlation with the criterion is the highest and its beta weight is consistently high and positive. Multiple correlations of the criterion with all predictors combined range from .53 to .69. The median R is .62. This R would account for approximately 38% of the variance between first year university average and a composite score of all predictors, which means approximately 62% of the variance would be attributable to accumulative error and other unknown factors.

Discussion

An attempt has been made to present the relationships between several intellective, non-intellective, and achievement oriented variables. In addition an attempt has been made to combine the variables in such a way as to predict most efficiently first year university average. Several points appear significant and are worthy of discussion.

Firstly, the correlations of the various variables with academic success at university are far below those reported in the manuals of these tests. Secondly, the predictive efficiency of all tests combined is of little practical utility. One reason for the low correlations could result from the low reliabilities of both predictor and criterion variables. It appears that after high school average is known little more is added to the effectiveness of the prediction of first year university average. For example, in Sample 6 the variance accounted for by high school average is about 36%. With the addition of eight variables the variance accounted for is about 48%, an addition of only 12%.

Examination of Table 2 reveals a disappointing though prescient fact. Many of the predictor variables fluctuate markedly in the magnitude of their relationship with the criterion. For example the CAAT Verbal score acts as a positive predictor within two samples, but as a

negative predictor within another. Such inconsistency is indicative of the distribution of the correlation coefficient for the predictor variable and as such is not totally unexpected. In any event it stands as a warning to minimize Type S errors (Lindquist, 1956) through careful subject randomization when conducting a study of this kind.

Table 2, Sample 6, shows the combination of predictors having the highest multiple correlation with the criterion. This sample also yielded the highest correlation between high school average and first year university average. By using only four variables from this sample a multiple correlation of .66 is obtained with the criterion. The variance accounted for is approximately 44%. The regression equation with the raw score beta weights of these four variables is as follows:

$y = -15.27 + .80\,x(1) + .92\,x(2) + .23\,x(3) + .17\,x(4)$ where
y = first year university average
$x(1)$ = high school average
$x(2)$ = Eysenck's Lie Scale
$x(3)$ = Attitude towards School (CSMS)
$x(4)$ = Verbal (CAAT)

Other lengthier equations can be obtained from Table 2, but the above equation was presented here because of its practical utility. That is, the tests in this equation can be administered in just over one hour making them more useful for educational counselling. It would seem desirable, too, that if tests are to be used for prediction purposes then regression equations like the one above should be employed. One should keep in mind, however, that there is a considerable amount of error involved in predicting scores when using such regression equations. On the basis of the present study prospects look dim for the usefulness of psychological tests in the prediction of freshman success at university.

References

Bendig, A. W. The Pittsburgh scales of social extraversion-introversion and emotionality. *Journal of Psychology,* 1962, 53, 199-209.

Brown, F. G., & Dubois, T. E. Correlates of academic success for high-ability freshman men. *Personnel and Guidance Journal,* 1964, 42, 603-607.

Costello, C. G. Two scales of need-achievement. In press.

Eysenck, S. B. M., & Eysenck, H. J. *The Eysenck Personality Inventory.* San Diego: Educational and Industrial Testing Service, 1963.

Kelly, F. J., Hunka, S., & Conklin, R. C. Further normative data on tests measuring flexibility in cognitive processes. *Psychological Reports,* 1965, 17, 683-686.

Lefcourt, H. M. Internal versus external control of reinforcement: A review. *Psychological Bulletin,* 1966, 65, 206-220.

Lindquist, E. F. *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mifflin, 1956.

Payne, R. W., Davidson, P. O., & Sloane, R. B. The prediction of academic success in university students: A pilot study. *Canadian Journal of Psychology,* 1966, 20, 52-63.

Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs,* No. 609, Vol. 80, 1, 1966.

Taylor, Janet A. A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology,* 1953, 48, 285-290.

Alberta Journal of Educational Research
Vol. XIV, No. 3, September, 1968

Techniques for directive counselling were developed on the basis of the Secord and Backman theory. This method of group counselling, when compared with a client-centered group counselling method, appeared more effective than the latter. Further experimentation with directive counselling seems justified on the basis of the findings.

JAMES W. BRIGGS
Edmonton, Alberta

A Comparison of Client-Centered VS Directive Group Counselling with High School Students

A major goal of counselling is the bringing about of some desirable change of behaviour on the part of those individuals counselled. By and large it is the troubled or inept or disturbed or unsuccessful student in high school who receives counselling. If counselling is to achieve its goal the student should behave after counselling in some different way that enables him to deal more effectively with his problems and to be a happier or more effective individual.

The manner in which a counsellor undertakes to help students bring about in themselves whatever changes in their own behaviour or outlook are necessary will depend upon the counsellor's theoretical orientation. To put it another way, if the counsellor is going to do more than simply fumble around in the hope that whatever he does will somehow benefit the student, he must proceed according to some plan and have some reasons for believing that whatever he is doing will work to the student's benefit. Each counsellor then must seek to adopt or develop a theory which seems to him to provide a reasonable rationale for his counselling practices. This is not an easy matter, however; first, because there are many theories of personality among which to choose and, second, because even when an acceptable theory has been chosen it is not a simple matter to learn how to tailor one's practice to fit the theory.

Faced with these hard facts there is a great temptation simply to follow a counselling procedure developed by some other practitioner (in line, it is presumed, with that practitioner's own theoretical bias). Although this should at least lead to some consistency in approach and a degree of confidence in what one is doing, the danger is that it may lead to a complacent attitude which inhibits sufficient attention to a realistic assessment of what results are actually being achieved.

No matter how convincing his rationale, each counsellor needs to face questions such as these: What is the evidence in my own experience that seeing students individually is more effective than seeing them in groups or *vice versa?* What is the evidence that one should counsel exclusively on either an individual or a group basis or by means of a combination of the two? What is my own evidence that a troubled student who is counselled does better in the end than one who is not counselled? (The record of proven effectiveness of talking therapy where attempts at evaluation have been made is somewhat disquieting! [Eysenck, 1952; Bergin, 1963].) What is the evidence that the theoretical basis for my own counselling methods is broad enough to include all the factors which are crucial, that no important aspects of behaviour or interpersonal relations are being ignored?

With these questions in mind let us begin by searching for a theory of personality broad enough to include as many as possible of the factors which influence an individual's behaviour. If such a theory were broad enough to encompass other theories of personality and capable of explaining both stability and change in an individual's personality, perhaps it could form the basis for some useful guide lines in the counselling process. The search for such a theory was the beginning of the present study.

Secord and Backman (1964) presented a useful synthesis of several theories of personality in a study reported in the *Psychological Review.* These authors began by trying to develop a theory which could account for both stability and change in an individual's personality and they were concerned with an interaction process which is truly interpersonal, where neither the behaviour of the individual nor of the other is the sole locus of cause.

In developing these ideas the notions of "interpersonal matrix" and "matrix congruency" are postulated. According to this view, the locus of behavioural stability and behavioural change is said to lie in the interpersonal matrix which has three components: Self-Concept, Self-Behaviour, and Other-Behaviour. There will be many matrices in the conceptual system of the individual. Matrices may be congruent or incongruent. A congruent matrix is achieved when all three components stand in a supportive relationship, i.e., "when the behaviors of the S (Self) and O (Others) imply definitions of self congruent with relevant aspects

of the self-concept" (Secord and Backman, 1961, p. 23). This sets out a well-structured framework and a very significant aspect of the idea of the dynamic interaction between the components.

Each component acts upon the other and changes are brought about in the components as a result of this. For example, the Self-Concept component interacts with the Self-Behaviour component and either or both of these components may be changed as a result of this interaction. Similarly, an interaction takes place between the Self-Components on the one hand and perceived Other-Behaviour on the other hand. But Other-Behaviour is itself modifiable by this interaction process and in turn, in the modified form, may have a different effect on Self. (Secord and Backman [1961, pp. 23, 24, 26, 27] cite some specific examples to illustrate this principle.)

These observations may be made clearer if the process is represented diagrammatically (Figure 1).

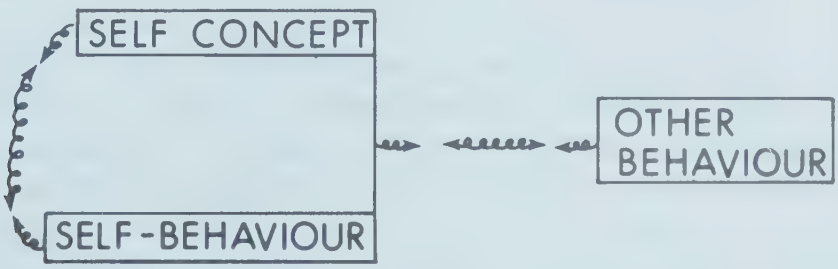


FIG. 1: A MATRIX IN THE PROCESS OF FORMATION

This diagram represents a matrix in the process of information where quite possibly one or more components are incongruent with the others. Since the tendency is always to seek congruency the interactions are probably bringing about changes in one or more of these components.

Eventually, at least in a relatively well-functioning individual, a state of congruency will likely be achieved. In its more stable form the personality matrix may now be represented as in Figure 2.

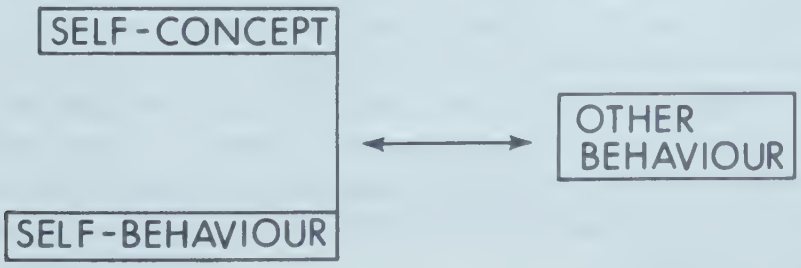


FIG. 2: REPRESENTATION OF CONGRUENT MATRIX

Here the arrows are smooth, representing a state of stability and smooth flow, but still double-headed to include the dynamic interaction factor

which is as essential a part of the matrix as any one of its components. As this situation occurs more and more often in more and more of the personality matrices or systems of matrices, the individual should function more and more smoothly and easily within his environment.

Looking at the structure of personality in this way, stability is seen not as something fixed or static in the usual sense, but as something depending upon a continuation of a smooth interaction process. A change, especially a sudden or marked change in any one component, could set the whole interaction chain within the matrix into oscillation. To restore a state of stability, changes may have to be made in one or both of the other components of the matrix. Thus, both stability and change are explainable in terms of this model. Furthermore, this point of view makes it possible to use many other theories of personality such as those developed by Harvey, Hunt and Schroder (1961), by Rogers (1951), by Festinger (1957), and by Kelley (1955) to explain, as it were, the working details of the processes indicated by the Secord and Backman model (Briggs, 1966).

If one can accept the point of view outlined above, then the next step is to consider the implications for counselling. If an individual is seriously disturbed or is functioning so poorly as to be seriously handicapped, the above point of view would lend one to conclude that he is finding himself unable to cope with the oscillations set up in the interaction process and is thereby unable to achieve sufficient congruency in the personality matrices. It may be assumed that extreme lack of congruency in many matrices would be very distressing.

In attempting to help such an individual the counsellor would need to keep in mind the triad of Self-Concept, Self-Behaviour, and Other-Behaviour and also the interaction among these three components. The aim of the counsellor would be to assist the counsellee to achieve a better measure of congruency in all three components, aware that he, the counsellor, is part of the client's Other-Component. The interaction process between the counsellor and the counsellee (which affects both of them) is one factor in achieving congruency. It may well be that the client needs to learn techniques that can be used to facilitate change in one or more of the components of the personality matrix. These techniques could very well include ways of influencing or changing others equally as much as techniques for changing his own Self-Concept or Self-Behaviour. Perhaps the counselling process should include some sort of "instruction" or "direction" with a view to helping the counsellee learn such techniques.

The Other-Component of the matrix may be very important, more so than many theories of personality seem to recognize. In any case there does not seem to be any reason to suppose that it is not equally important as the Self-Concept or Self-Behaviour components. If this is so, it is

possible that group counselling is much more necessary than has been realized. The individual counselling relationship may have very real limitations since the Other-Component is limited and rather drastically modified. It is possible that group counselling where more than one "other" is present in the therapeutic environment is a very necessary part of a properly balanced counselling procedure. This is in contrast to one widely followed counselling method (usually but not always carried out on an individual rather than group basis) which concentrates deliberately on Self-Concept.

In brief, the implications for counselling seem to be that while attentions to the Self-Concept of the individual should certainly be part of the counsellor's concern, he should not neglect or ignore the other components of the personality matrix nor should he ignore the interaction processes that seem to go on among the components. There is a suggestion that a necessary part of the counselling process may be a directive sort of relationship aimed at helping the counsellee to acquire techniques of handling the interaction processes so that he gains more facility in bringing about a congruent state among all three components of his personality matrices. Thus, group counselling in itself may not be sufficient since it is entirely possible to use a client-centered approach in a group situation where the chief concern is with the Self-Concept of the individuals within the group.

One might now hypothesize, in line with the Secord and Backman theory, that counselling which concerned itself with all three of the matrix components and also with the interaction process would be more effective than a purely client-centered approach which is preoccupied with the Self-Concept component. To test this out one might work with two groups in one of which a client-centered approach alone was used while with the other a directive approach fitting in with the Secord and Backman theory was used. This, in essence, is what was attempted in the present study in which three groups of troubled individuals were compared on the amount of improvement over a given period of time. Three treatments were compared: no counselling, group counselling by a client-centered method, and group counselling by a more directive method in which attention was given to all three components of the personality matrix and to the techniques of handling effectively the interaction processes. The specific hypotheses tested were:

1. The uncounselled group would show the no improvement.
2. The directed group would show more improvement than the client-centered group.

Fifty students in an urban high school were selected for study. For each student there appeared to be reasonably conclusive subjective evidence that he was experiencing more than usual difficulty in interpersonal relationships. From among these fifty students three groups of five

girls and three groups of six boys were chosen randomly. Each group in each sex category was then randomly designated either as control (no counselling), client-centered (that is, slated to receive client-centered counselling) or directed (that is, slated to receive directed counselling as described below). Before the counselling began, all students were asked to complete the I.P.A.T. Junior-Senior High School Personality Questionnaire (Cattell, 1962). At the conclusion of the experimental period the students were again asked to fill in the same questionnaire and a comparison of scores was made using t tests of the significance of differences between group means.

The control groups were not counselled. The client-centered groups (one group of boys and one group of girls) were counselled as a group for five sessions of thirty-five minutes duration not more than once a week over a period of eight weeks. The client-centered method of counselling is well-known to most counsellors and those not familiar with it can find many good descriptions of this method, notably Rogers (1951). The remaining groups were counselled for similar periods and the same number of times using a directed approach based on the Secord and Backman theory.

In the "directed" approach the total series of counselling sessions was deliberately planned and a "structure" was maintained by the counsellor who directed the attention of the counsellees to aspects of behaviour and of cognitive processes that were felt to be crucial in learning to build more congruent systems of personality. The chief concerns of the counsellor were these:

1. To see that a definite, pre-arranged progression was followed as the interview sessions went on so that each component of the matrix received attention.

2. To see that attention was paid to the techniques of dealing with the interaction processes since "According to the interpersonal congruency theory, the individual actively uses techniques or mechanisms for maintaining his interpersonal environment so as to maximize congruency" [Secord and Backman, p. 584]. The structure of this series of counselling sessions consisted of three phases.

Phase One: The Self-Concept

The aim of the early counselling sessions was to direct the attention of the individuals to their own Self-Concept in order to clarify it and make it as explicit as possible. The procedure was to help them reveal their own values and ideals, to struggle to reach tentative conclusions in areas of doubt, to accept differences in one another's ideas and ideals and to accept the idea that one's set of values is to some extent subject to change and modification as experiences accumulate.

Phase Two: The Self-Concept—Self-Behaviour Interaction

Here discussions were initiated on how to act in a manner consistent with one's own ideas of right and wrong, to identify sound bases on which to make such decisions, to decide what to do about errors and mistakes. The subject of semantics was brought in on a non-technical level using ideas by such writers as Johnson (1946); especially useful was Johnson's idea of the error of dichotomizing "success" and "failure."

Phase Three: The Dynamics of Self-Other Interaction

Here attention was directed to such matters as genuine consideration of others, selective interaction with others, empathy as a skill to develop, development of the ability to observe and really listen to others. It is not suggested that philosophical or technical discussions took place during this phase but relations with others were discussed in practical terms with a view to developing perception and a realistic appreciation of others.

A question will naturally arise as to whether or not the counsellor did indeed use two different and clearly distinguishable methods of counselling with the two experimental groups. To check this out, tapes were made of the first five counselling sessions for each of the client-centered and the directed groups. These ten tapes were then played in random order to a group of eight competent judges each of whom was asked to indicate independently which type of approach was used in each taped session. For all the directed sessions there was 100% agreement that this was the method used. For three of the client-centered sessions there was 100% agreement that this method was used while for the remaining two 75% of the judges agreed that this was the method being used.

The results of the analysis of the change between first and second scores on sub-tests of the I.P.A.T. Questionnaire were compared. Statistically significant differences were found for only four sub-tests: E—submissive vs dominance, G—weak vs strong superego strength, J—given to group action vs obstructive individualism, and O—self-assured, placid vs apprehensive, worrying. These results are reported in Table 1.

An examination of the data gives some support to the following statements:

1. The only significant change noted for the uncounselled groups was in the direction of increased worry and apprehension. This supports the first hypothesis that uncounselled groups would show no improvement.

2. One client-centered group showed a significant change in the direction of dominance as opposed to submissiveness. Trends were also noted among the client-centered groups indicating an increase in sen-

TABLE 1

SIGNIFICANT CHANGES BETWEEN FIRST AND SECOND TESTINGS[a]

Factor	Uncounselled	Client-Centered	Directed
E		girls (.025)	boys (.025)
G			boys (.10)
J			girls (.10)
O	boys (.10)		

[a]Probability given in parentheses.

sitivity, an increase in self-assurance and a decrease in worry and apprehension. This gives a degree of support to the hypothesis that client-centered counselling would result in some improvement.

3. One directed group showed a significant increase in dominance as opposed to submissiveness and also a significant trend toward being less rule-bound. Another directed group showed a significant change toward more ready acceptance of group action as opposed to obstructive individualism. A trend was noted in these groups toward increased self-concept control and self-discipline. These findings lend some support to the hypothesis that directed counselling would bring about the most improvement.

It seems worthwhile to note the following trends even though they cannot form the basis for any firm conclusions because significance of difference was not obtained. There was a consistent trend in the results obtained on second testing, compared to the first test scores, which seemed to indicate that directed counselling may be more effective than client-centered counselling in increasing self-concept control and self-discipline. Client-centered counselling seemed to have little positive effect here and may even have had an inhibiting effect since both control (uncounselled) groups made more gains on this factor than the client-centered groups. It might be conjectured, though it was by no means proven, that attention to the techniques of bringing about congruence in the total personality matrix (which was attempted with the directed groups) may have resulted in more successful control of self-concept and self-behaviour. Such a result if it can be established in subsequent experimentation would be consistent with the Secord and Backman theory.

In conclusion it must be said that the significant results were too few to allow any categorical statement as to the superiority of one counselling method over the other. In general there was some vindication of group counselling *per se* inasmuch as the uncounselled groups not only showed no significant improvement but in one instance showed significant regression.

In vindication of directed counselling based upon the Secord and Backman theory, one can say that it appeared to be at least as effective as client-centered counselling and may have been somewhat superior in the beneficial results obtained. There was at least sufficient evidence to justify further experimentation with this type of counselling. Perhaps the evidence will be thought sufficient to encourage other counsellors to familiarize themselves with the Secord and Backman theory and to carry on further experimentation along the lines suggested here. If there is even a reasonable possibility that a directed counselling procedure based upon this theory could provide more help per counselling hour for our high school students, it is worth further investigation. In the uncertain state of psychotherapy today, any new procedure which offers hope for improvement ought not to be ignored.

References

Bergin, A. E. Effects of psychotherapy: Negative results revisited. *Journal of Counseling Psychology,* Fall, 1963, 10, 244-250.

Briggs, J. W. A comparative study of client-centered versus directive group counselling. Unpublished Masters thesis, The University of Alberta, 1966.

Cattell, Raymond B., & Beloff, Halla. *Handbook for junior-senior high school personality questionnaire.* Institute for Personality and Ability Testing, Champaign, III., 1962.

Eysenck, H. J. The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology,* 1952, 16, 319-324.

Festinger, L. *A theory of cognitive dissonance.* Stanford University Press, 1957.

Harvey, O. J., Hunt, D. E., & Schroder, H. M. *Conceptual systems and personality organization.* New York: John Wiley & Sons, Inc., 1961.

Johnson, Wendell. *People in quandries.* New York: Harper & Bros., 1946.

Kelley, George A. *A theory of personality.* W. W. Norton & Co., Inc., 1955.

Rogers, Carl R. *Client-centered therapy.* Boston: Houghton Mifflin Co., 1951.

Secord, Paul F., & Backman, Carl W. Personality theory and the problem of stability and change in individual behavior. *Psychological Review,* 1961, 68, 1, 21-32.

Secord, Paul F., & Backman, Carl W. *Social Psychology.* New York: McGraw Hill Book Co., 1964. P. 584.

Alberta Journal of Educational Research
Vol. XIV, No. 3, September, 1968

There is a close relationship between social class and membership on the high school student council. Council members come from a higher social class. They are also more active in extra-curricular activities, but not necessarily sports activities. These same students feel more politically efficacious in the area of student decision-making, and hold a higher sense of citizen duty toward student elections.

A. R. MacLEOD
Ponoka, Alberta
and
W. D. KNILL
The University of Alberta

Students' Council Leadership

Every high school will have a unique set of aims for the students' council and its system of student government. However, it is generally agreed that the program of student government must contribute to the effectiveness of the school. Both the school and the individual must profit by the activities resulting from the students' council program. According to Van Pool (1964), the single most important objective is that of "teaching habits of responsible citizenship." By examining a number of factors which may be associated with council leadership, this study (MacLeod, 1966) sought to determine whether the aim of responsible citizenship and other commonly accepted aims of student government are being realized.

Theoretical Framework

An initial factor important in determining who will be elected to the student council is the presence or absence of school regulations which restrict membership for some or all of the offices. Knill (1963) found that:

> The relationship between grade level and the proportion of students holding elected positions is something that is built into the system by tradition (i.e. students in the lower grades are not expected to be candidates for as

> many offices as are the senior students) or by rules and administrative control (i.e. students in the lower grades are not permitted to run for top positions in the council) [p. 61].

There have been a number of studies in the area of extracurricular activities, and to a greater extent in the social sciences, that suggest the importance of three variables in determining who holds office: social class, attitudes toward the political process, and membership in voluntary organizations.

Hollingshead's (1949) study of *Elmtown's Youth,* tested the hypothesis that "The social behavior of adolescents is related functionally to the positions their families occupy in the social structure of the community [p. 9]." An examination of the Elmtown students' union revealed that students from the upper classes were disproportionately represented on council while the lower class students did not elect as many representatives as their numbers in the student body warranted.

The general area of attitudes toward the political process has been studied by Campbell (1952) and Presthus (1964) while Ziblatt (1965) and Knill (1963) examined students' attitudes toward politics. Knill used a modification of Campbell's Political Efficacy Scale and Sense of Citizen Duty Scale to obtain student evaluation of the council as an agent for their political welfare, of themselves as effective agents in council affairs, and of themselves in the role of voters. He found that girls felt more efficacious than boys as voters in student council elections and that students felt increasingly more efficacious as they progressed to the higher grades.

The relationship between membership on the student council and participation in voluntary organizations was found to be consistent with the findings of Presthus (1964). For the communities of "Edgewood" and "Riverview" a pronounced relationship between class status and membership in voluntary organizations was noted. This was the same in the ten high schools studied by Knill (1963).

Research Design

The study was designed to determine whether selected variables are related to the election of an individual to a students' council. The sample surveyed consisted of the grade ten, eleven, and twelve students in attendance during the 1965-66 school year in the Grande Prairie School District. Data gathered by questionnaire were used to examine the social class of council and non-council students, to ascertain the extent to which members of the student council participate in extra-curricular activities as compared with the overall student population, and, finally, to compare the attitudes of council members toward the political process, as reflected in student government, with that of the other students.

The instrument used in collecting the data was devised especially for this investigation. The first section of the questionnaire consisted of items designed to measure student attitudes toward the political process. Questions taken from Campbell's study of voting were modified for administration to high school students. The Guttman technique for scalogram analysis was used to determine whether the items formed a unidimensional scale. When items scale, the scale score an individual is assigned provides a description of his attitude with respect to each of the scale items. Further, as persons obtaining the same score have responded similarly, then they can be expected to be similar in the characteristics which the responses might represent. A second section of the questionnaire elicited demographic and general information such as age, grade, sex, council membership, membership in voluntary organizations, and occupation of the father. In subsequent analysis the Blishen Scale (1958) was used to translate the father's occupation into a social class category.

Results

The first hypothesis tested was that the social class of student council members would be higher than that of non-members. Social class was determined by father's occupation according to the Blishen Scale. The resulting frequency distribution is presented in Table 1. Means and standard deviations were computed for social classes of council and non-council members. A "t" test of the significance of the difference of means between the two groups indicated a "t" significant at the .0001 level (two tailed). As might be expected, the application of a Mann-Whitney "U" test to the same data resulted in a "U" significant at the .0002 level (two tailed). The above results strongly support the first hypothesis that Council and Non-Council Members differ in social class, and an inspection of the means indicates that the difference favours Council Members (see Table 2).

The second hypothesis stated that "council members express a more favorable attitude toward the political process as tested by the scale of political efficacy, the scale of sense of citizen duty and the scale of contempt for people." The students responded to the following forced —choice statements:

Political Efficacy Scale:

Voting is the only way that students like me can have any say about how the students' council runs things.

Sometimes students' council activities and business seem so complicated that a student like me can't really understand what's going on.

Students like me don't have any say about what the students' council does.

I don't think student council members care much what the students like me think.

TABLE 1

FREQUENCY DISTRIBUTION OF SOCIAL CLASS

BY COUNCIL MEMBERSHIP

	Social Class Category							
Council Membership	1	2	3	4	5	6	7	Total
Member	1	20	4	3	7	7	1	43
Non-Member	15	39	26	15	131	48	29	303
Total	16	59	30	18	138	55	30	346

TABLE 2

TEST OF SIGNIFICANCE OF THE DIFFERENCES BETWEEN SOCIAL

CLASS MEANS OF COUNCIL AND NON-COUNCIL MEMBERS

	Mean	Standard Deviation	t
Council Members	3.47	1.70	4.03*
Non-Council Members	4.55	1.64	

* Significant $p < .0001$

Sense of Citizen Duty Scale:

So many other students vote in the students' council elections that it doesn't matter much to me whether I vote or not.

It isn't so important to vote when you know your candidate doesn't have any chance to win in the students' council election.

A good many students' council elections aren't important enough to bother with.

If a person doesn't care how an election comes out he shouldn't vote in it.

The Political Efficacy Scale produced a coefficient of reproducibility of 92.8, and the Sense of Citizen Duty Scale produced a coefficient of reproducibility of 94.2, both above the criterion level of 90.0. Differences in scale scores were all in the hypothesized direction, but only the Political Efficacy Scale differentiated strongly between council and non-council members. Table 3 presents the results of Mann-Whitney "U" tests.

TABLE 3

MANN-WHITNEY "U" TESTS OF DIFFERENCES BETWEEN

COUNCIL AND NON-COUNCIL MEMBERS

Scale	U	Z	P
Political Efficacy	5338.0	-2.684	.0037
Sense of Citizen Duty	6820.5	-0.350	.3632

An item by item analysis of the scales was undertaken to obtain more detailed information about student attitudes toward the political process. Utilizing both "t" tests and "U" tests it was found that the item which provided the greatest differentiation between the two groups was from the Political Efficacy Scale. This item, significant at the .0001 level, was "Students like me don't have any say about what the student council does." Attitudes were then examined for sex differences but of the twenty questions used to test attitudes only one reached significance at the .05 level.

The third hypothesis stated that "there is a greater proportion of council members who participate in voluntary organizations than there is of non-members." Three questions were designed to test this hypothesis and two further items considered the relation between executive membership in voluntary organizations other than the student council and council membership.

Students were classified by council membership and by participation in extracurricular activities. A chi-square test showed significance at the .01 level indicating that a significantly greater proportion of council members were participants in extracurricular activities than was the case for non-members. A subsidiary hypothesis to the effect that council members are those who participate in three or more extracurricular activities was supported by a chi-square test (significant at .01 level).

One might expect that there would be more executive members of clubs, both within the school and outside the school, that would be council members than would be the case for the rest of the student group. However, there was little relation between non-school executive membership and council membership. A chi-square test produced a result that could be expected to occur by chance more than fifty times out of a hundred. In the case of school clubs over one-half the council members were on club executives while only one-seventh of the non-council students held such positions. This relationship was statistically significant at the .001 level.

Finally, the common view that the student who participates in interschool sports will be elected to a council position was not strongly supported. There were about 10 chances in 100 that the obtained chi-square could have been obtained by chance.

Implications

There is a need for research on students' councils that is based on theories which are not merely statements of what some individual believes ought to be, but rather are descriptive theories which can give rise to empirical predictions that can be confirmed in experience. Unfortunately, until educators feel a need to examine critically what they are accomplishing when they set up systems of student government, they will be prescribing functions which they desire but will remain oblivious to the actual functioning of the council as an institution furthering democratic principles. Thus, although many educators firmly state that students' councils are providing training for all students in the operation of democracy, in actuality, as was indicated in the present study, the council may be a narrowly constituted body which is not representative of the student body as a whole. The feeling of a large group of students that they cannot influence the political process may partially account for a feeling of apathy toward the council and its activities.

References

Blishen, B. The construction and use of an occupational class scale. *Canadian Journal of Economics and Political Science,* 1958, 4, 156-172.

Campbell, A., *et. al. The voter decides.* Evanston: Row, Peterson, 1952.

Hollingshead, A. B. *Elmtown's youth.* New York: Wiley, 1949.

Knill, W. D. The adolescent society of the high school. *The Canadian secondary school.* Toronto: Gage and MacMillan, 1963.

MacLeod, A. D. Factors related to election to a students' council executive. Unpublished M.Ed. thesis, The University of Alberta, 1966.

Presthus, R. *Men at the top.* New York: Oxford University Press, 1964.

Van Pool, G. M. The student council. *The Bulletin,* 1964, 48, 42-53.

Ziblatt, D. High school extracurricular activities and political socialization. *The Annals,* 1965, 361, 20-31.

BOOK REVIEW

CONCEPTUAL FOUNDATIONS OF SCIENTIFIC THOUGHT

By: Marx W. Wartofsky (New York: The Macmillan Company, 1968).

I find reviewing this recent contribution to the literature of philosophy and the philosophy of science a pleasure. In his preface the author stated that "the main emphasis throughout is on the philosophy of science as a philosophical discipline." This is an accurate description of the form and content of the book and consequently the chief reason for its being, in the opinion of this reviewer, a most significant contribution to the field. Wartofsky has achieved what most writers in the philosophy of science have failed to achieve. He has written a philosophical treatise in the context of science and scientific thought that is neither pseudo-philosophy nor pseudo-science. For this, he is to be congratulated and it may be said that he has contributed significantly to bringing the philosophy of science to philosophical maturity. It may be said too, and argued with vigour, that the book deserves to be classified as scientific epistemology rather than as philosophy of science. I doubt if Wartofsky would recognize or accept this distinction because the emphasis he displays shows that he may view philosophy of science and epistemology as synonymous fields. This is a view with which I would not wish to disagree, and the exemplification of which I find encouraging.

The book contains four parts, I "The Genesis of Scientific Thought," II "The Methods of Science," III "Some Fundamental Concepts in the Sciences," and, IV "Coda."

Only one chapter—"Science, Values, and the Humanistic Understanding"—is included in the final part. It is a short chapter and serves only to relate scientific inquiry to the humanistic domain. One could wish that the author had developed this dimension more fully, but as considerable attention is given to it incidentally throughout the book the brief note in the final chapter may be justified.

The book also contains four appendices including a short description of logical notation, which should prove very useful to the less philosophically sophisticated reader, and a very comprehensive set of bibliographical notes. These latter make the book a valuable source for further information, particularly as the entries are clearly and accurately annotated.

Part I is in content and style somewhat reminiscent of the writings of Karl Popper—particularly his *Conjectures and Refutations* (London: Routledge & Kegan Paul, 1963) but this reviewer found Wartofsky's exposition somewhat more readable. It should prove interesting and

beneficial to students in philosophy and particularly to those in science. Those in the latter category are usually so grossly naive with respect to the conceptual bases of the processes and procedures to which they are expected to subscribe, that it is hoped a book such as this will be recommended to their attention. Throughout this part Wartofsky, unlike Popper, is content to maintain a descriptive and explicative role and to avoid propounding a particular point of view. He advances a view of science as human activity, describes its development from the pre-scientific era and explicates the contributions to its genesis by the Greeks with particular emphasis on science as critical activity. The result is a coherent exposition that makes foolish the slavish adherence to the "Two Cultures" attitude popularized by C. P. Snow. In fact, Wartofsky's book could be sub-titled "A Refutation of the Two Cultures Myth."

The six chapters in Part II are devoted to a philosophical exposition and analysis of scientific methodology from the problems of "observation" through "models," "measurement," "experiment," "induction and probability" to the context of "explanation" by laws and theories. This is perhaps the most rewarding part of the book. The author assiduously maintains a philosophical perspective by giving the essential minimum of the methodological characteristics necessary to make his philosophical probings relevant and comprehensible. In our age which is typified by glib statements about and references to such things as "models," "experiments," "tests," "laws," and "theories" it is refreshing to find a responsible statement about them that sets each in a relevant perspective. This reviewer has long been interested in and concerned by the inadequate conceptual foundation of educational research. For his students, Wartofsky's book is now a required study.

In Part III Wartofsky treats "Causality," "Space, Time, and Matter," "Reduction and Explanation in the Biological Sciences," and "Some Fundamental Concepts in the Human Sciences." The four chapters cohere and must be viewed as a whole. If read as a whole, they relate the scientific and the humanistic as two interrelated and integrated aspects of a single culture. Wartofsky's is not a stereotyped exposition of the mechanistic processes nor is it a mere description of a series of elements. It is a probe into the conceptual foundations and therefore consistent with the aim and the title of the book.

Marx Wartofsky has written a good book which can be confidently recommended to philosophers, scientists in all domains, and students in all fields. It is readable and should not prove too difficult for a significant proportion of laymen, and those who are willing to devote some effort to its study are not likely to remain appropriately categorized as laymen.

G. R. Eastwood
Simon Fraser University

NOTES TO CONTRIBUTORS

FREDERIC J. BOERSMA is an Associate Professor in the Department of Educational Psychology at The University of Alberta. He received his Ph.D. from Michigan State University in 1965. His current research interest is in the relation of eye movements and pupillary responses to concept formation. He has had articles published in many well known journals, among them *Psychological Review, Journal of Educational Psychology, Educational and Psychological Measurement*, and the *Journal of General Psychology.*

JAMES W. BRIGGS, who received his Master's degree in Educational Psychology from The University of Alberta, was formerly a high school counsellor and an assistant principal in the Edmonton Public Schools. He is now the principal of Glenrose School Hospital, a school for the rehabilitation of multiply handicapped children, K-12, in Edmonton.

RODNEY C. CONKLIN holds a joint appointment with Student Counselling Services and the Department of Educational Psychology at The University of Calgary. He received his M.Ed. from The University of Alberta and is working towards his Ph.D. at the same university. His publications, in the area of personality measurement and learning, have appeared in *Perceptual and Motor Skills, Psychological Reports, Psychology in Schools*, and *Psychonomic Science.*

GORDON R. EASTWOOD graduated B.A., M.Ed., from the University of Western Australia. He was awarded a Fulbright Scholarship and earned the M.A. and Ph.D. degrees from the University of Minnesota. He has held positions at the University of Toronto, and New England, N.S.W., Australia and is currently Head, Department of Social and Philosophical Foundations at Simon Frazer University. He has published articles in professional journals in Australia, Canada and the United States. His primary interest is in the application of the philosophy of natural and social science to problems encountered in educational research.

HENRY W. HODYSH received his doctoral degree from the Department of Educational Foundations at The University of Alberta. He is now an Assistant Professor in that department.

W. D. KNILL is a Professor of Educational Administration at The University of Alberta. He is presently on sabbatical and will teach in

Australia on exchange for part of the year. In 1964, he won the EdPress Award for a series of articles on value of orientations of high school students.

ARTHUR G. KNOWLES is an Assistant Professor at the University of Victoria. He received his Ph.D. in Educational Psychology at The University of Alberta. He has had articles published in *The Alberta Journal of Educational Research.*

A. R. MacLEOD obtained his B.Ed. and B.A. in 1963 from The University of Alberta. In 1966-67 with the aid of Graduate Research assistant ship and Alberta Teachers' Association Scholarship he completed his M.Ed. in Educational Administration. At present he is the Assistant Principal at the Ponoka Composite High School, Ponoka, Alberta.

DONALD G. OGSTON received his M.Ed. from the Department of Educational Psychology, The University of Calgary. He is now holding a one year appointment in Student Counselling Services at that university.

B. PATTERSON received his M.Ed. in Educational Psychology from the University of Calgary. He is now a guidance counsellor for the Calgary Public School Board.

ARTHUR G. STOREY, Ph.D. Stanford, who is Associate Professor of Educational Psychology, The University of Calgary, has published more than thirty research papers. He also received the Ryerson Press Fiction Award for his novel *Prairie Harvest* (1959), and has subsequently published a number of short stories in Canada and England.

The Alberta Journal of Educational Research

Vol. XIV, No. 4 *December, 1968*

CONTENTS

BOOK REVIEWS

COMMITTEE ON EDUCATIONAL RESEARCH
The University of Alberta

Editorial Comment

The final issue of Volume XIV of the Journal includes articles on a range of topics. There are two articles on concept formation in young children, an article considering variables related to the social status of high school students, a discussion of the relationships among measures on male elementary education students, and an article dealing with the interaction among members of school staffs. This range of topics illustrates the broad responsibilities of educators who must constantly be concerned with the education of children and youth, and with those who undertake the task of teaching.

Everett and Armstrong present an interesting study of the responses of young children to concrete and pictorial materials. Their careful analysis, taking account of the interaction among the variables, leads to more definite implications concerning the selection and use of concrete and pictorial representations in helping the young to form concepts.

Liedtke and Nelson approach the question of concept formation in young children from yet another point of view. Liedtke, who is fluent in two languages and has an elementary reading ability in a third, was concerned about the apparent evidence that the bilingual child is handicapped in concept formation. The findings of this study, if interpreted within the limitations as set out by the authors, may add another dimension to the study of bilingualism and concept development.

Chabassol and Thomas present the findings of an investigation of the relationships among anxiety and performance in teacher education for male students enrolled in elementary education. Studies of sub-groups in which it is difficult to get the prized "large N" are rarely reported in educational research. These authors report a companion study to their investigation of female elementary education students which was reported in an earlier issue of the Journal.

The study by Zentner and Parr includes an examination of the variables which seem to be related to high social status among high school students. These results seem to be incompatible in some respects with the results of similar studies of American high school students.

Miklos and Brietkreuz discuss the use of an interaction model for examining the structural characteristics of schools. The model most frequently used to analyze these characteristics is the bureaucratic model. The authors conclude that the analysis of communication and interaction patterns within school staffs may be used with other than small samples and that the data from such an analysis are a useful supplement to information gathered in a study of the bureaucratic dimensions of schools.

P.A.L.

Alberta Journal of Educational Research
Vol. XIV, No. 4, December, 1968

The study compared young children's responses to concrete and pictorial modes of representing perceptual data. Concrete materials were more effective than pictorial materials, with age significantly affecting the level of response. Sex and socio-economic status had no significant effect on level of response.

LORENE M. EVERETT
and
ROBERT D. ARMSTRONG
The University of Alberta

Responses of Young Children to Two Modes of Representation[1]

The relationships among language, thought, and reality can be exemplified most clearly in concept formation. Working with concrete perceptual data, the young child first identifies then labels the stimuli of his environment. Later, he assembles categories and names them, building a rich fabric of meaning around the concepts and their verbal tags. The child's definition of a word suggests the level reached in this search for more concentrated, i.e., more abstract, meaning. Thus, cognitive development matches linguistic growth, paralleled by a movement to increasing use of more abstract stimulus materials until pictorial representations and even words can be manipulated, categorized, and defined. The processes by which the child uses reality to think and to produce ever more sophisticated language are not well known but can be observed to some degree through a study of language. The fluency, precision, and abstractness of his language should be an index of the power of the child's thought.

It is, then, well known that the very young child first uses concrete materials as a source for perceptual data and that, increasingly, more abstract materials are substituted. It was the purpose of the present study to determine whether concrete or pictorial representations would evoke more meaningful responses from young children. It was hypothesized that more meaningful responses would be elicited by the use

[1] Based on an unpublished M.Ed. thesis by Lorene M. Everett, Department of Elementary Education, The University of Alberta.

of concrete stimuli and that the quality of the responses would be conditioned by age, sex, and socio-economic status.

Sample

The sample of sixty children consisted of ten boys and ten girls at ages three, four, and five, with half of the children at each age level coming from low socio-economic status (LSES) homes and half from high socio-economic status (HSES) homes. Socio-economic status was determined by combining the occupational rating of both parents as listed on the *Blishen Occupational Scale* (1961). Children from the two extremes of the socio-economic rating scale were selected from the study, with those considered HSES coming from homes where the parents were employed in professional occupations, many having university or highly specialized training, while parents of those considered LSES were employed in unskilled occupations. All subjects were attending preschool centres in Edmonton, Canada.

Test Procedure

Eighteen objects represented in two different modes were: shoe, saw, apple, cow, turtle, stove, dress, airplane, scissors, bed, bread, train, corn, elephant, truck, hammer, jeans, and chair (Figure 1). The two modes of representation were (*a*) concrete (toy representations of familiar objects), and (*b*) pictorial (black and white line drawings).

In order to assess the ability of the children to respond to the two modes, three tasks were administered individually to each child, first in one mode of representation and three weeks later in the other mode, with the order of presentation being varied systematically. Responses were recorded on a dictaphone, transcribed on cards by a qualified stenographer, and scored by the experimenter. Reliability of scoring was ascertained by interscorer agreement.

As a means of establishing rapport with each subject, several stimulus objects, in whichever mode was being used in the test situation, were displayed, manipulated, and talked about by the subject and the investigator prior to Task 1.

For Task 1 (Definition) the child was presented with a stimulus cue (either concrete or pictorial), asked to name it, then encouraged to tell as much about it as he could. Responses to Task 1 were scored according to the modified *Feifel-Lorge Qualitative Scale* (1950) which was revised on the basis of results of a pilot study which had utilized several rating scales as a means for evaluating children's definitional responses (Everett, 1968). Each response was scored according to the following scale:

(a) Synonym Category (e.g. orange—a fruit) =3 points,
(b) Explanation, Use, Description Category (e.g. orange—you eat it; it grows on trees) =2 points,

Fig. 1. Examples of stimulus materials in concrete and pictorial modes of representation.

(c) Demonstration, Repetition, Illustration and Inferior Explanation (e.g. puddle—a puddle of water; eye—point to eye) =1 point.

A score was obtained for each cue stimulus by giving credit for all responses elicited by that stimulus.

A count of all the scorable responses was also recorded in order to determine what, if any, relationship there was between quality and quantity of responses.

In Task 2 (Categorization), twelve stimulus cues were placed on a table, the child was asked to label each object, then six more stimuli were presented, one at a time, with the child being asked in each case to look at all the pictures (or toys) that were on the table and find the ones that were the same as or like the stimulus. The subject was encouraged to look carefully at all stimuli and be quite certain that he had included all those he thought belonged together.

Upon indicating that he had finished, the subject was asked to explain the reason he had grouped together the items he had selected.

The subject received credit in Task 2 (*a*) for the number of items placed in a logical category, and (*b*) for the explanation given for the grouping according to *Annett's Scheme of Classification* (1959) as follows:

(a) Class name (e.g. animals, tools) =4 points,
(b) Similarities (e.g. corn, apple, bread—eat all of them) =3 points,
(c) Contiguity (e.g. corn, cow—the cow eats the corn) =2 points,
(c) Enumeration (e.g. the cow is on a farm, the turtle is in the water) =1 point.

In Task 3 (Identification), six cue stimuli were placed upon a table. Three definitions in the form of questions had been prepared for each of the six cue stimuli, based on a hierarchy of levels of conceptualization from abstract to concrete. The experimenter then presented the child with the highest level definition for each of the cue stimuli, one at a time, asking the child to indicate which of the six stimuli on the table had been defined. The child was free to identify the correct stimulus by naming, pointing to, or using any means that indicated an understanding of the appropriate object.

The subjects were able to identify each stimulus at one of three levels ranging from concrete to abstract, with a score of three obtained for identification at the most abstract level; two points at the functional level; and one at the concrete level.

Results and Discussion

Responses to the three tasks were subjected to three-way analyses of variance with repeated measures on the last factor. The .05 level of significance was used to evaluate the findings. Age, mode of representation, socio-economic status and sex were the variables considered. A summary of the findings follows.[2]

[2] Complete information on the statistical analyses may be obtained from the authors.

1. Age: On each of the tasks, age was found to have a statistically significant effect, with an increase in age resulting in an increase in level of achievement. Upon examining mean scores it was discovered that there were substantial gains on quality of definitions given; from age three to four (concrete—gain of 6 points; pictorial—gain of 8 points), and from four to five (concrete—7; pictorial—8). A smaller increase was found in quantity of scorable responses between three and four years (concrete—3; pictorial—3) and between four and five years (concrete—4; pictorial—4). These results seem to indicate that although the quantity of responses increases steadily with age, the quality increases more rapidly, resulting in superior definitions given in fewer words as children get older (Figure 2).

The gains made from one age level to the next on the categorization

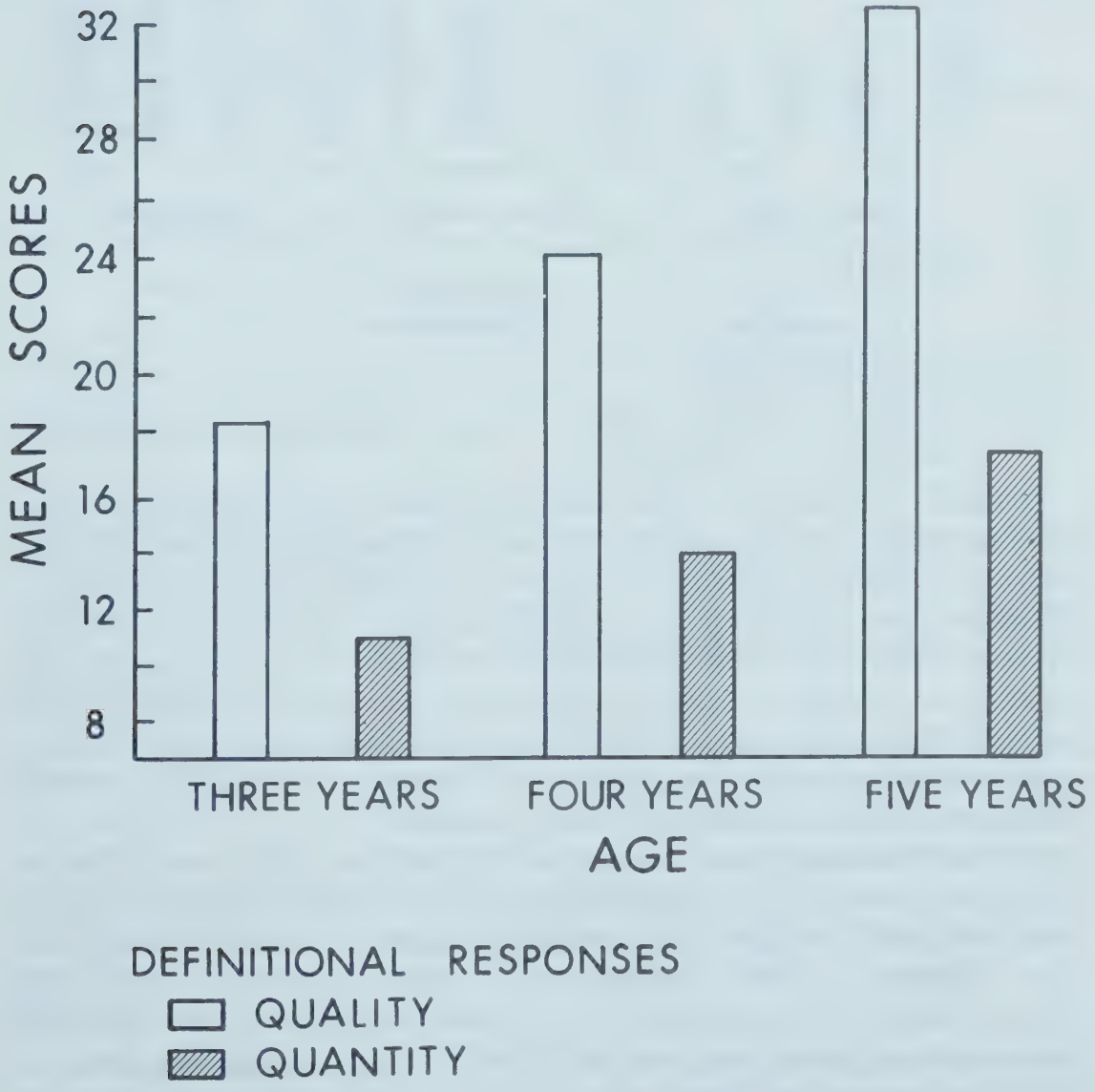


Fig. 2. Comparison of mean scores for age groups on quality and quantity of definitional responses.

task were also significant, with the ability to produce and explain groupings at a higher level of conceptualization increasing with age.

Similarily, but to a lesser degree, the ability to identify at a more abstract level increased with age (Figure 3).

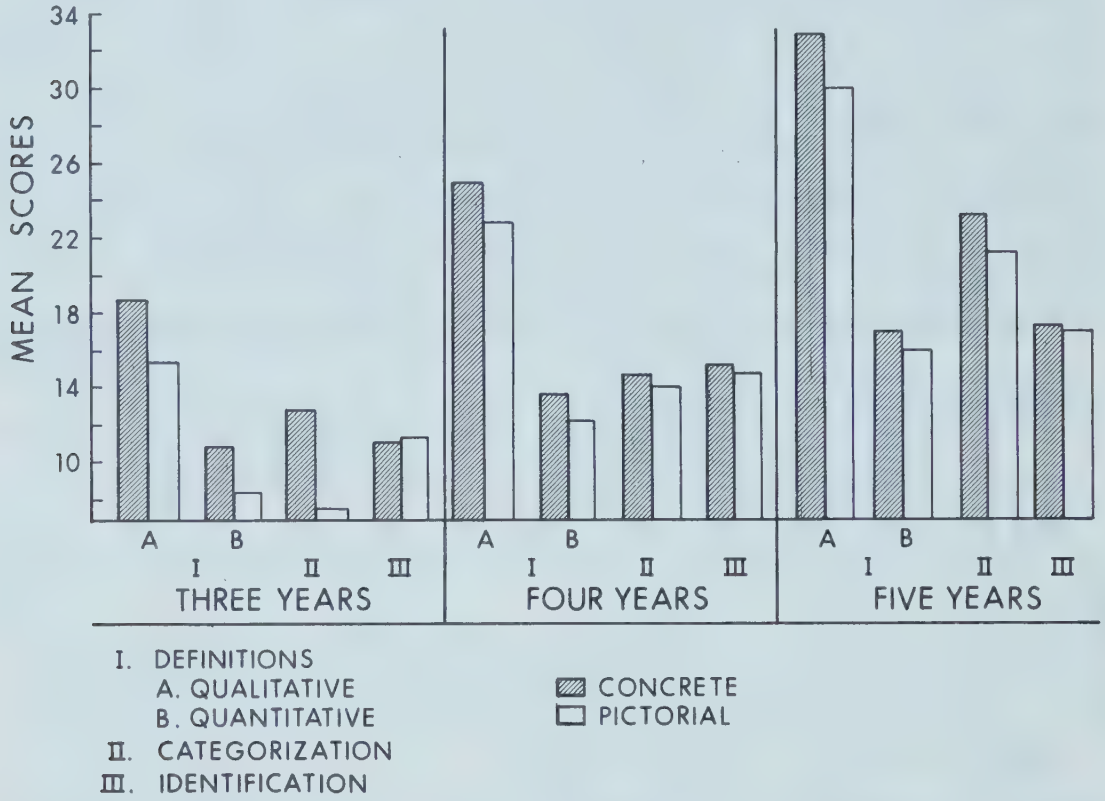


Fig. 3. Comparison of mean scores for age groups on test tasks for concrete and pictorial modes of representation.

2. Mode: In Task 1 (Definition) and Task 2 (Categorization), the concrete mode of representation evoked significantly superior responses compared to those evoked by the pictorial mode. However, on Task 3 (Identification), mode of stimulus presentation resulted in little difference in the quality of responses, with the three year old children actually performing at a slightly higher level in response to the pictorial mode. Although the concrete mode evoked superior responses on Task 1 and 2 at the four and five year levels, the differences in response between modes was not so great as for three year olds.

3. Socio-Economic Status: Studies utilizing the *Feifel-Lorge Qualitative Scale* (Grant, 1965; Labensohn, 1967) have clearly shown that verbal ability of HSES children is signifciantly superior to that of LSES children. However, in the present study, when the child's conceptual understanding was measured by non-verbal, as well as verbal means, socio-economic status was not found to have any significant effect on ability to perform in either mode.

4. Sex: No sex differences were found in ability to deal with either mode of representation on any of the three tasks.

Conclusions and Implications

As concrete materials were found to elicit greater conceptual understanding in young children than did pictorial materials, it would seem important that teachers of preschool and primary children recognize this difference in stimulus potential. Young children's difficulties in dealing with mathematical and language symbols may arise from a lack of ability to deal with pictorial stimuli. Thus, it would seem important that numerous experiences with concrete objects are provided in early childhood education programs.

Even though most of the subjects could correctly label all of the stimuli in both modes, this did not assure that they could talk meaningfully about or classify the cue stimuli. Therefore, if a measure of a young child's conceptual understanding is desired, it would seem likely that a test which required the child to carry out a variety of tasks in both concrete and pictorial material would be most valid. Obviously, minimal recognition as a measure of vocabulary knowledge would seem inadequate.

The difference between performance in response to the two modes of representation was much greater for three year olds than for the five year olds, which seems to indicate that three year olds are more tied to concrete sensory attributes than are five year olds. Early childhood educators should devise programs to help children see the relationship between various modes of representation and aid them in making a transition from the subjective, sensor-motor phase to a more objective, abstract phase of thinking.

If the trend displayed by the three, four, and five year old children described in the study reported here continues into primary grades, it would seem that primary grade children should be quite competent in dealing with familiar concepts in either mode. However, new concepts might be more meaningfully developed through the employment of concrete material.

Contrary to the findings of a similar study by Sigel, Anderson, and Shapiro (1966), it was found that ability to deal with phenomena in varying modes of representation was not influenced significantly by socio-economic status. However, this may be due to the lack of extremes in socio-economic status of the people living in Edmonton. A similar study with larger numbers of children from extremely disadvantaged areas might indicate a significant difference in their ability to deal with the two modes of representation.

Since some trends were indicated here, a longitudinal study including a greater span of years might serve to confirm present indications and, also, to determine the influence of such variables as sex and socio-economic status.

References

Annett, M. The classification of instances of four common class concepts by children and adults. *British Journal of Educational Psychology,* 1959, 29, 223-236.

Blishen, B. R. The construction and use of an occupational class scale. In B. Blishen, et. al. (Eds.) *Canadian Society.* Toronto: The Macmillan Company of Canada Limited, 1961. Pp. 477-486.

Everett, L. M. M. Responses of young children to two modes of representation. Unpublished Master's Thesis, The University of Alberta, 1968.

Feifel, H., & Lorge, I. Qualitative differences in the vocabulary responses of children. *Journal of Educational Psychology,* 1950, 61, 1-19.

Grant, M. A qualitative analysis of the vocabulary responses of good readers and poor readers. Unpublished Master's Thesis, The University of Alberta, 1965.

Labensohn, D. The effects of socio-economic status on recognition and recall vocabulary of grade one children. Unpublished Master's Thesis, The University of Alberta, 1967.

Sigel, I., Anderson, L., & Shapiro, H. Categorization behavior of lower and middle class Negro preschool children: Differences in dealing with representation of familiar objects. *Journal of Negro Education,* 1966, 35, 218-229.

Alberta Journal of Educational Research
Vol. XIV, No. 4, December, 1968

Two samples of Grade One pupils, one monolingual and the other bilingual, were tested on a specially constructed Concepts of Linear Measurement Test based on Piaget's test items. The bilingual sample proved to be significantly superior to the monolingual sample on the concept formation test.

WERNER W. LIEDTKE
and
L. DOYAL NELSON
The University of Alberta

Concept Formation and Bilingualism

Introduction

Research reports on bilingualism are numerous, and most investigators have dealt with the relation of bilingualism and intelligence. In one of the more recent studies (Peal and Lambert, 1962) it was reported that bilingualism has favourable intellectual consequences. The study conducted by these authors yielded the following results: in a comparison of bilingual and monolingual children, the bilinguals were intellectually superior, possessed greater verbal skills, exhibited greater mental flexibility, were more facile at concept formation, and achieved higher grades in school. It was the purpose of the study reported here (Liedtke, 1968) to consider just one of these stated results and test for the effects of bilingualism on concept formation.

Concept Formation

Perhaps the greatest quantity of work in the field of concept formation has been conducted by the Swiss psychologist Jean Piaget. He has concerned himself not only with the learner, but also with the substance to be learned and its structure or logical organization. With his ingenuous methods and experiments he has probed into the growth of children's concepts in a variety of areas. In his books on number (1952) and geometry (1960) Piaget makes brilliant suggestions as to how concepts develop out of simple behaviour patterns, and how a basic repertoire of numerical and geometric concepts is formed. His

investigations help us realize that children must assimilate more and more experience in order to make valid generalizations later. The results of his experiments vividly demonstrate that the development of mathematical concepts and ideas is a long and slow process.

Concept formation represents a major part of intellectual development. According to Piaget's theory, intellectual development proceeds through qualitatively distinct stages that occur in a definite sequence. During each stage the child acquires the abilities that prepare him for the next stage. The rate of preparation, according to Piaget, is determined by four factors: maturation, experience, social interaction, and equilibration. Experience and social interaction are cultural rather than genetic factors. There is an implication that the sequence of stages is fixed, but that the ages at which they occur is not. However, Piaget's developmental approach to behaviour change "is characteristically not concerned with any systematic exploration of other independent variables which may temporarily accelerate or retard the appearance of the behavior studied [Flavell, 1963, p. 17]."

Experience and social interaction are the two main factors responsible for individual differences or retardation and acceleration of concept formation. Adler (1966) in commenting on proposed revisions of the arithmetic program concludes that acceleration of learning is not ruled out by Piaget's theory of stages in the development of thinking. At least two of the contributing factors, the experience acquired in interaction with the physical environment and the influence of the social milieu, are subject to our control since they are cultural rather than genetic.

Bilingualism

Many children are brought up in a bilingual home environment. They learn two languages simultaneously at a very early age. Thus, they are exposed to an environment that is in a way unique and complex. Most of them have two 'worlds' of experience. The factor of social interaction in Piaget's developmental theory includes the explicit and implicit teaching of the child by other people in his environment. Learning two languages or two symbols for every object implies a great degree of explicit teaching. One might hypothesize that a young child learning two languages at the same time is exposed to a greater amount of social interaction when compared to someone his own age learning just one language. The question arises whether the bilingual child's experiences are different enough to effect his intellectual development in any significant way.

As mentioned previously, there exist numerous studies that have been concerned with bilingualism. Jensen (1962) in a two article summary lists many disadvantages of bilingualism ranging from speech

handicap and intellectual immaturity to emotional instability and social maladjustment. However, he points out that most of these results came from studies that used varying definitions of bilingualism, few subjects, and a great variety of procedures, and argues that bilingualism may indeed be an advantage to the child's intellectual development.

Peal and Lambert (1962) present a thorough review of studies on bilingualism. They found that most of the reports support the detrimental effect of bilingualism on intellectual development, a few found no significant difference between monolinguals and bilinguals, and only two studies concluded that bilingualism has favourable intellectual consequences. Also stated in their report is the fact that many of the studies failed to define monolingualism and bilingualism, did not control for socioeconomic status differences, compared different age groups or failed to consider the educational background of their subjects. A carefully controlled study by the investigators contradicted much of the existing evidence about bilinguals. In comparing the intelligence of bilinguals and monolinguals they found that on the tests administered the bilinguals scored significantly higher on all verbal and most non-verbal parts of their tests. On none of the subtests did the monolinguals score higher than the bilinguals. Most bilinguals achieved higher grades in school. The bilinguals possessed greater verbal skills, exhibited greater mental flexibility and they were more facile at concept formation.

The Study

Purpose and Hypotheses

It was the purpose of this study to consider the experience of becoming bilingual at any early age and to test for its effects on mental development. A comparison of certain aspects of concept development of bilingual and monolingual children was then made. To make the comparison a test on concepts of linear measurement was constructed to serve as the main instrument. The test consisted of six subtests which dealt with the following aspects of linear measurement: (*a*) reconstructing relations of distance, (*b*) conservation of length, (*c*) conservation of length with change of position, (*d*) conservation of length with distortion of shape, (*e*) measurement of length, and (*f*) subdividing a straight line. Each of the sub-tests has items similar to items devised by Jean Piaget (1960, pp. 67-149). The subtests were used to find out how children judge distances and how they come to an understanding of conservation and measurement of length.

The results of the *Concepts of Linear Measurement Test* were used to test the following hypotheses:

When adjustments are made for differences in age, kindergarten attendance, sex, intelligence, and socioeconomic status:

(*a*) Monolingual and bilingual children do not differ in their understandings of linear measurement concepts (Total Test).

(b) Monolingual and bilingual children do not differ in their ability to conserve length (Subtests 1-4).
(c) Monoglingual and bilingual children do not differ in their ability to measure length (Subtests 5 & 6).

The data gathered on age, kindergarten attendance, sex, intelligence, and socioeconomic status were also used to test whether any one of these was a significant factor in predicting scores on the *Concepts of Linear Measurement Test.*

The Samples

The two samples required for this study came from six schools of the Edmonton Separate School System. Nine grade one classrooms were made available.

The criterion for choosing the monolingual subjects was that they had no functional knowledge of a second language and that only one language was spoken at home. One hundred thirty-seven pupils from six of the classrooms met this criterion. A table of random numbers was used to select fifty subjects. The resulting sample consisted of twenty-five girls and twenty-five boys.

Bilinguals were defined as children who had used two languages before entering school and who were exposed to both languages at home. The classroom teachers helped in selecting the fifty subjects for the second sample from the remaining three rooms which were set up for bilingual children. The bilinguals received instruction in English and French, whereas the monolinguals were instructed solely in English. There were twenty-five females and twenty-five males in the bilingual sample.

Data Collected

The *Concepts of Linear Measurement Test* was administered individually to each subject during the month of May, 1968.

Each subject's age was calculated to the nearest month on the day of testing. The socioeconomic status for each subject was established by referring to the Occupational Class Scale constructed by Blishen (1961). A numerical value was given to each subject according to the father's occupation. All the subjects had written the S.R.A.—Primary Mental Abilities Test (Thurstone, revised 1962, Grades K to 1) a month prior to the investigation.

Results

To find out whether the two samples differed significantly on the basis of age, socioeconomic status and intelligence, *t*-ratios were calculated. The results of the *t*-tests are shown in Table 1.

A check was made on each subject's pre-school experience. Fifty-two per cent of the monolinguals and seventy-four per cent of the

TABLE 1

T-TESTS COMPARING MEANS OF THE

SAMPLES ON AGE, SEX, AND INTELLIGENCE

	Means		
	Monolinguals	Bilinguals	t^a (calculated)
Age (years)	6.85	6.82	0.435
Socioeconomic Status	49.16	49.56	0.265
Intelligence	107.40	104.22	1.325

[a] t (.05) = 1.985

bilingual sample had attended either kindergarten or playschool. Since the difference in attendance for the two samples was rather large, the significance of kindergarten attendance in predicting scores on the *Concepts of Linear Measurement Test* was checked. Independent of the group membership, the kindergarten attendance and the scores on this test were correlated. The resultant correlation coefficient of 0.141 was not significant at the .05 level.

The test on concepts of linear measurement was scored and the means, variances, and standard deviations for the two samples and the total group were calculated. The results are shown in Table 2.

TABLE 2

MEANS, VARIANCES AND STANDARD DEVIATIONS OF

'CONCEPTS OF LINEAR MEASUREMENT TEST' SCORES FOR

MONOLINGUALS, BILINGUALS, AND TOTAL SAMPLE

	Mean	Variance	Standard Deviation
Monolinguals	10.72	13.88	3.72
Bilinguals	13.72	13.72	3.70
Total Group	12.22	15.76	3.97

Multiple linear regression analysis was used to analyze the results in terms of the research hypotheses. The results of the analysis are summarized in Tables 3 and 4.

TABLE 3

MULTIPLE-LINEAR REGRESSION ANALYSIS OF THE HYPOTHESES

Source	Degrees of Freedom Num.	Den.	F-Ratio
Total Test	1	94	16.04[a]
Part 1 (Conservation)	1	94	8.41[a]
Part 2 (Measurement)	1	94	18.08[a]

[a] $F_{.05}(1,94) = 3.95$

TABLE 4

RESULTS OF ANALYSIS FOR INDEPENDENT VARIABLES

Source	Degrees of Freedom Num.	Den.	F-Ratio
Age	1	99	0.73
Kindergarten Attendance	1	99	2.02
Sex	1	99	2.87
Intelligence	1	99	4.16[a]
Socioeconomic Status	1	99	3.85

[a] $F_{.05}(1,99) = 3.94$

Conclusions and Discussion

On the basis of testing the research hypotheses the following conclusions may be drawn:

(*a*) The mean for the bilingual sample on the *Concepts of Linear Measurement Test* was significantly higher than the mean for the mono-

lingual sample. This result is in agreement with Peal and Lambert's (1962) finding that bilingualism has favourable effects on intellectual functioning. From their study they concluded that bilinguals were more facile at concept formation.

(*b*) The mean for the bilingual sample on the 'Conservation' part (Part I: subtests one to four) of the test was significantly higher than the mean for the monolingual sample. If the higher score implies that the concept is more advanced and more highly developed. the bilingual children manifest a better understanding of the concept when compared with monolingual children of the same age.

(*c*) The mean for the bilingual sample on the 'Measurement' part (Part II: subtests five and six) of the test was significantly higher than the mean for the monolingual sample. This suggests that the measurement concept, too, had developed to a more advanced stage in the bilingual subjects.

These results seem to indicate that the linguistic and cultural experience of the bilinguals is an advantage. The evidence would seem to demonstrate the importance of social interaction and social environment as ingredients of experience. Intelligence factors necessary for concept formation seem to be developed to a greater extent in the bilingual subjects.

Of all the independent variables considered only intelligence (I.Q. scores) was found to be a significant factor in predicting scores on the *Concepts of Linear Measurement Test*. This result is in agreement with Almy (1966) who reviewed several studies that related intelligence and results of Piagetian tests of concept formation. She concluded that the safest generalization one can make is that to some degree "brightness" pays off. Pelletier (1966) also found that children with high mental ability were definitely more advanced in their development of the concepts of linear measurement than children with low mental ability.

It is probable that the chronological age range was not large enough and the socioeconomic levels represented here were not diverse enough to produce any significant correlation with the test score.

The results of this study give rise to several implications. If bilingualism increases intellectual potential and is beneficial to concept formation then a second language should be introduced during the early years when experience and environmental factors are most effective in contributing to the development of intelligence.

The results of this study seem to indicate that becoming bilingual speeds up the normal process of some parts of mental development. According to Peal and Lambert (1962) the speaking of two languages diversifies the intellectual structure. Teachers and administrators who deal with or instruct students who are bilingual should consider that

these pupils could be advanced in the development of certain concepts and make provisions for curricular adjustments which would take such development into account.

The results seem to indicate that bilingual children conserve length before monolinguals do. If this is true for other conservations, it could be that bilingualism accelerates development, and the bilinguals reach the concrete operational stage before the monolinguals do. This of course would have important implications for those who teach such students in the primary grades.

References

Adler, I. The Cambridge report: Blueprint or fantasy? *The Arithmetic Teacher,* 1966 (March), Vol. 13, 179-186.

Almy, M., Chittenden, E., & Miller, P. *Young children's thinking.* New York: Teachers College Press, Teachers College, Columbia University, 1966.

Flavell, J. H. *Developmental psychology of Jean Piaget.* New York: D. Van Nostrand, 1963.

Jensen, J. V. Effects of childhood bilingualism. *Elementary English,* 1962, Vol. 39, 132-149, 358-366.

Liedtke, W. W. Linear measurement concepts of bilingual and monolingual children. Unpublished Master's Thesis, The University of Alberta, Edmonton, 1968.

Peal, E., & Lambert, W. E. The relations of bilingualism to intelligence. *Psychological Monographs,* 1962, Vol. 76, No. 27, 1-23.

Pelletier, J. D. A study of grade one children's concepts of linear measurements. Unpublished Master's Thesis, The University of Alberta, Edmonton, 1966.

Piaget, J. *The child's conception of number.* London: Routledge & Kegan, 1952.

Alberta Journal of Educational Research
Vol. XIV, No. 4, December, 1968

A summary of the intercorrelations obtained for scores in anxiety, academic aptitude, academic achievement and teaching performance for 50 male elementary education students. Negative correlations were obtained between scores for anxiety and for all other variables. A comparison with an earlier study based on female students reveals a number of marked differences.

DAVID J. CHABASSOL and DAVID C. THOMAS
University of Victoria

Anxiety, Aptitude, Achievement and Performance in Male Elementary Teachers

The purpose of the present study was to determine the relationship between the non-intellectual factor, anxiety, and the variables academic aptitude, scholastic achievement, and performance in a given area, to wit, classroom teaching. This study is essentially a replication of an earlier study done by the same authors (1967) which utilized a population of female teachers.

Method

Scores on the Taylor *Manifest Anxiety Scale* (MAS) provided a measure of anxiety (Taylor, 1953). Academic aptitude was evaluated by the *School and College Ability Test* (SCAT), and final grade point averages (GPA) served as measures of academic achievement. The subjects used in this study were 50 male elementary student teachers and scores made on the above variables were correlated with final practice teaching grades (FPTG) in each case.

The subjects were in their second year at the University of Victoria and were enrolled in the Faculty of Education. This second year is the professional year for such students and, with the exception of a second year English course, all work is taken in the Faculty of Education. The year previously the students had taken five courses, four from the Faculty of Arts and Science and one, a first year course in educational psychology, from the Faculty of Education.

In their professional year these students are required to teach for a total of about eight weeks. At the end of the practice teaching sessions, a final grade in teaching is calculated for each student teacher. This grade is a composite of the evaluations from eight to twelve cooperating teachers and faculty members who have observed the student's teaching and have provided criticism and advice. Grade point averages are also calculated for each student at the end of the April examination period. This grade point average for professional year students is influenced by the mark received for practice teaching. The latter contributes 33 1/3% to the former, so that a spuriously high positive correlation might be expected between the two scores.

The *Manifest Anxiety Scale* was given to the students in their first term. To the original fifty items of the MAS were added an additional one hundred items not intended to measure anxiety, but designed to mask the true nature of the scale. Taylor's items were arranged in three groups, in the order suggested by Hoyt and Magoon (1954). Those items which Hoyt and Magoon found to be most closely correlated with clinical assessments of anxiety were grouped so as to yield the score MAS_1. The items which were next in agreement with clinical assessments of anxiety yielded the MAS_2 score. Those items poorest in this respect yielded the MAS_3 score. The total scores, MAS_t, for all three subtests combined, were also obtained for each subject. All students involved in this study had taken the *School and College Ability Test* on being enrolled at the University of Victoria in their first year and three scores, verbal ($SCAT_v$), quantitative ($SCAT_q$), and total ($SCAT_t$)) were available.

Results

Table 1 provides the following information:

1. All subtest scores on the MAS correlated significantly with each other and with the total MAS score. It will be seen that MAS_3, calculated from those items which Hoyt and Magoon found to be the poorest indicators of clinical anxiety, correlate somewhat more poorly with MAS_1 and MAS_2 than do these subtest scores with each other, and also less well with MAS_t than do MAS_1 and MAS_2.
2. The twelve correlations between measures of anxiety, MAS, and measures of academic aptitude, SCAT, are evenly divided into positive and negative results. However, whereas none of the positive correlations approach significance, two of the negative ones do, though only at the .05 and .10 levels.
3. $SCAT_v$ and $SCAT_q$ scores are positively related, with significance beyond the .001 level.

TABLE 1

CORRELATIONS BETWEEN MAS, SCAT, GPA, AND FPTG SCORES FOR MALE ELEMENTARY TEACHERS

	MAS_1	MAS_2	MAS_3	MAS_t	$SCAT_v$	$SCAT_q$	$SCAT_t$	GPA	FPTG
MAS_1	---	.68*	.55*	.89*	.14	.09	-.32#	-.09	-.30#
MAS_2	---	---	.45*	.84*	-.09	-.02	-.12	-.11	-.19
MAS_3	---	---	---	.79*	.22	.13	-.19	-.12	-.22
MAS_t	---	---	---	---	.11	.08	-.25†	-.13	-.28#
$SCAT_v$	---	---	---	---	---	.73*	-.25†	.00	-.21
$SCAT_q$	---	---	---	---	---	---	-.13	.11	-.10
$SCAT_t$	---	---	---	---	---	---	---	.22	.88*
GPA	---	---	---	---	---	---	---	---	.60*
FPTG	---	---	---	---	---	---	---	---	---

*p = < .001 #p = < .05 †p = < .10

4. $SCAT_v$ and $SCAT_t$ scores are negatively related, but with significance beyond the .10 level of confidence only.
5. $SCAT_q$ and $SCAT_t$ scores are negatively related, but the finding is not a significant one.
6. All correlations between MAS scores and GPA are negatively related, but not significantly so.
7. SCAT scores tend to be positively correlated with GPA, but the results do not approach significance.
8. All MAS scores are negatively correlated with FPTF. In two cases the correlation is beyond the .05 level of confidence.
9. $SCAT_v$ and $SCAT_q$ are negatively correlated with FPTG, but the findings are not significant.
10. $SCAT_t$ and FPTG are positively correlated, with significance beyond the .001 level of confidence.
11. GPA and FPTG scores reveal a correlation significant at the .001 level but, as noted above, this relationship is spuriously high.

Discussion

The high and positive correlations between the various subtests of the MAS with each other and with MAS_t were to be expected. However, the correlations between $SCAT_v$ and $SCAT_q$ and the relation each of these bears to $SCAT_t$ comes as a surprise. One would not have expected a correlation of the magnitude of .73 between the verbal and quantitative scores of the SCAT, as these subtests supposedly measure quite different abilities. Further, inasmuch as the $SCAT_v$ and $SCAT_q$ scores contribute to the $SCAT_t$ score, it is surprising to find negative scores between the two subtest scores and the total score. The relationships between all MAS scores and SCAT scores are in the expected direction if we assume that high scores in anxiety would interfere with, or depress, measures of academic ability. A similar, negative relationship is noted between anxiety scores and measures of academic achievement. Again, the direction of the correlation is as we might expect but the magnitude of the findings is not impressive. A similar comment might be made with reference to the relationship between all SCAT scores and GPA. One would imagine that measured academic ability would bear a much closer relationship with final grades in courses taken than is revealed here. Final grades in teaching are negatively related to high scores in anxiety, as might be expected. It is difficult, if not impossible, to explain the correlations noted between SCAT and FPTG scores. Whereas $SCAT_v$ and $SCAT_q$ reveals negative relationships with teaching grades, the total SCAT score shows a very high and positive correlation with final grades in teaching. Indeed, for this group of elementary male teachers, the $SCAT_t$ score is the single best predictor of teaching success.

A comparison of the results of this study with those of the earlier study by Chabassol and Thomas which utilized 124 female elementary teachers is of interest. There are a number of quite remarkable changes in both direction (i.e., sign) and in magnitude of the correlations noted in these two studies. It was found in the present project that the twelve correlations between MAS and SCAT scores were evenly divided between positive and negative results, with two of the negative results significant at the .10 level. In the study involving females, eleven of the twelve correlations were negative, although none was significant. Relationships between SCAT subtest scores and between these scores and $SCAT_t$ differ remarkably in the two studies. The correlation between $SCAT_v$ and $SCAT_q$ was .15 for females, but .73 for males. Again, the correlation between $SCAT_v$ and $SCAT_t$ was .82 for females and -.25 for males, with the former being significant beyond the .001 level of confidence, positively, and the latter significant beyond the .10 level of confidence, in the negative direction. Finally, $SCAT_q$ and $SCAT_t$ scores reveal correlations of .65 for females but -.13 for the present male group. Of the four MAS and FPTG correlations in the earlier study, three were

positive, one was negative, and none was significant. In the present project, all these correlations were negative, and two were significantly so. The most noticeable difference between the two sets of findings relates to the correlation between $SCAT_t$ and FPTG. For the female population the correlation was found to be -.21, but for the present group it was seen to be .88 and, as noted above, the single best predictor for teaching success for males. On the basis of the above comparisons, it can be stated that correlations between SCAT - MAS, SCAT - FPTG, MAS - FPTG, and $SCAT_t$ - FPTG differ quite notiveably, according to the sex of the subjects involved. Those who would predict success in teaching on the basis of scores made on other variables must take into account the sex of the teachers for whom the predictions are being made.

References

Chabassol, D., & Thomas, D. Anxiety, aptitude, achievement and performance in female teachers. *Alberta Journal of Educational Research,* 1967, Vol. XIII, No. 4, 291-294.

Hoyt, D. P., & Magoon, T. M. A validation study of the manifest anxiety scale. *Journal of Clinical Psychology,* 1954, Vol X, 357-361.

Taylor, J. A. A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology,* 1953, Vol. XLVIII, 285-290.

Alberta Journal of Educational Research
Vol. XIV, No. 4, December, 1968

Influence and social structures in school organizations are investigated through an analysis of communication and interaction patterns among staff members. It was concluded that this method of analysis need not be confined to case studies nor to small schools, and that is useful as a supplement to methods based on the bureaucratic model.

E. MIKLOS
The University of Alberta

ELMER A. BREITKREUZ
Devon, Alberta

Analysis of Influence and Social Structures in Schools

Most recent approaches to the analysis of the structural characteristics of schools viewed as organizations have relied heavily upon descriptions which involve bureaucratic or similar dimensions (MacKay, 1964, Kolesar, 1967). Variations in structural characteristics are described in terms of variations in measures obtained from the information which members of the organization provide about the presence or absence of specific characteristics. This research strategy has proven to be useful in yielding insights into the nature of school and school system organization; however, like any single approach, it also suffers from deficiencies as far as providing a complete and accurate portrayal of the object under study.

In the first place, such research tends to focus on a fairly limited aspect of structure, namely on those features most closely related to the formal or deliberately planned parts of the organization. It does not even begin to provide information about the complex network of relationships which develops within the formal framework; that is, it does not offer adequate means for identifying and analyzing significant centres of power and influence or the channels of communication through which influence is exerted. Secondly, in spite of attempts to allow for the interrelatedness of parts and the dynamic character of an organization, research based on the bureaucratic model is struggling essentially with a static

model and presents significant but limited possibilities for contributing to the development of a more comprehensive model of organizations. The research approaches outlined in this paper are not presented as alternatives to the type of research just described. Instead these strategies are proposed because of their possibilities for supplementing and complementing studies of bureaucratic characteristics to provide a more complete description of an organization and expanded possibilities for testing hypotheses about organizational dynamics.

Problem and Research Procedures

The research project described in this paper was based on the assumption that useful information about significant structural characteristics of school organizations can be obtained through the analysis of communication and interaction patterns which develop among members of the school staff.[1] Specifically, the project was designed to study the influence and subgroup structures in relation to the following questions:

a. What are the characteristics of school influence structures, and what are the bases of personal influence?
b. What are the characteristics of task and non-task related subgroup structures?
c. How much overlap is there between task and non-task related structures?

Similar studies have been carried out by Boyan (1951), McCleary (1957), and House (1966) among others; however, in each of these instances the objective of the study was to provide an intensive analysis of a single school. The present study was designed to examine similarities and differences across a sample of schools. Although questions and hypotheses were suggested by previous research, the study was not based on any clearly delimited theorizing about the characteristics of influence and subgroup structures. It will become evident in the report that the study was guided mainly by attempts to test some commonly held notions about the formation of subgroups within larger organizations and about the bases of interpersonal influence.

Sample

The sample for this study consisted of 18 non-city schools whose staff members were willing to participate in the research project. The schools varied in size from 14 to 33 staff members; a variety of types in terms of grade range, from elementary to senior high school, was included. Members of the research team visited all but one of the schools for the purpose of administering the questionnaires; in the one instance responses were obtained by mail. Completed questionnaires were obtained from 389 staff members in the 18 schools, including principals, vice-principals, counsellors, teachers, secretaries, and librarians.

[1] The research project was supported by a grant from the Alberta Advisory Committee for Educational Studies.

Instrumentation

The instruments and research procedures used in this study were based upon and were extensions of those described by Blocker, McCabe, and Prendergast (1964). Basic data were provided by the responses of staff members to six sociometric-type questions. Respondents circled the code numbers of all those persons with whom they interacted or communicated in six difference contexts; that is, the data provided six "slices" of interaction patterns in the schools at a particular point in time. The six dimensions: communication, socialization, three reliance dimensions, and attributed influence were defined operationally by these questions:

a. During the course of a typical school week, in school or out of school, with which individuals are you likely to discuss general school matters (teaching duties, school events, etc.)? (General Task Related Communication)
b. With which individuals are you most likely to socialize informally during recesses, during noon hours, and/or before and after school hours? (Socialization)
c. If you had a problem concerning discipline in your classroom from whom would you likely seek advice? (Reliance—Discipline)
d. If you had a problem concerning the organization of teaching materials, teaching methods, tests or assignments, from whom would you likely seek advice? (Reliance—Teaching)
e. If you had a problem concerning the interpretation of school policies and regulations from whom would you likely seek advice? (Reliance—Policies)
f. In your opinion, which individuals in this school are most influential in initiating changes in general school practices such as testing programs, school regulations, school activities, etc.? (Attributed—Influence)

Responses to all but the second question were used in the analysis of influence structures; the first two questions yielded the data for the analysis of subgroup structures.

Analysis

The basic procedures which have been developed for the analysis of sociomatrices were applied in this study (Glanzer & Blaser, 1959; Blocker, McCabe & Prendergast, 1964).[2] A sociomatrix was constructed for each question for each school; a *one* was entered in a cell to indicate that communication or interaction had taken place between individuals and a *zero* was entered if no interaction was reported. This first power matrix included all of the primary communication, socialization, reliance, and attributed influence links. For further analysis, only reciprocated links were retained in the first two matrices; these matrices were then squared and cubed to reveal two and three-step channels of communication (Blocker & McCabe, 1964).

2 Techniques specific to this study are presented in Blocker. Computer programs were prepared by Lawrence Bezeau now with the Department of Elementary Education, The University of Alberta.

Influentials on the communication dimension were determined from column sums in the cubed matrix; these sums revealed the total number of three-step communication links between an individual and all other persons in the school. Column totals were used to rank order staff members and the upper one-fifth of staff members was classed as influential in that school while others were classed as non-influential. An examination of this categorization for all schools indicated that this arbitary procedure provided a reasonably good separation between those individuals who had larger and those with fewer numbers of tertiary communication links.

In order to determine influentials on the reliance dimensions, a subweight substitution method suggested by Blocker and McCabe (1964) was used. Column totals for the first power matrix were considered to be subweights for each individual. These subweights were substituted for non-zero elements in the third power matrix; the columns for this matrix were then summed and each individual's subweight was added to his column total. Staff members were again placed in rank order and those in the upper fifth in each staff were classed as influentials.

Since the attributed influence matrix did not lend itself to expansion as did the others, members were simply ranked on column totals.

The detection of cliques or subgroups has persisted as one of the most intriguing problems in sociometric analysis. Although the inspection of sociograms, the manipulation of rows and columns of the first power matrix, the manipulation of higher power matrices have all proven useful to a point,[3] it is only since the attempts of Wright and Evitts (1951) and MacRae (1960) to factor analyze sociometric data that subgroup detection has been possible for larger groups or organizations. For the purpose of detecting subgroups in the socialization and communication dimensions, the factor analysis procedures suggested by Blocker and his associates were applied and extended (Blocker, McCabe & Prendergast, 1964).

The intercorrelation matrix derived from the third power communication or socialization dimension matrix was subjected to principal axis factor analysis, and iteration was continued until eigenvalues below 1.00 were reached. Through the analysis of results and the examination of independently prepared sociograms it appeared useful to accept a factor loading as low as 0.40 on a varimax rotation as indicative of subgroup membership. Factor analysis results consistently verified independently prepared sociograms for larger schools and contributed greatly to the completion of this task.

Results

Influence Structure

The nature of the influence structures in the schools under study was analyzed in relation to the following questions: Do teachers classed

[3] A review of techniques for the analysis of subgroups is presented in Glanzer and Blaser, pp. 324-328.

as influentials differ from non-influentials on selected professional and personal characteristics? Do persons who are influential on one dimension also tend to be influential on the others? Where are administrators located within the influence structure? Table 1 summarizes the results of t-tests and Chi-square tests used to test hypotheses of no significant differences between influentials and non-influentials.

TABLE 1

SUMMARY OF SIGNIFICANCE OF DIFFERENCE TESTS BETWEEN INFLUENTIALS AND NON-INFLUENTIALS ON SIX CHARACTERISTICS[a]

Influence Dimension	N[b]	Age	Total Exp.	School Exp.	Years Prep.	Grade Level	Sex
General Communication	77 312	0.16	1.47	2.76**	3.34**	2.38*	10.68**
Reliance-Discipline	65 324	1.15	4.49**	2.70**	7.20**	5.28*	49.21*
Reliance-Methods	71 318	2.64**	4.28*	3.50*	6.84**	4.58**	27.16**
Reliance-Policies	67 322	1.89*	3.28**	2.69**	4.15**	2.59*	28.49**
Attributed Influence	69 320	0.57	2.98**	2.30*	7.21**	4.88**	41.38**
Generalized Influentials	40 349	2.47**	4.14**	2.95**	6.54**	4.39**	--

[a] t-test except for Chi-square on sex differences.

[b] Upper value = number of influentials; lower value = number of non-influentials

*p ≤ .05; **p ≤ .01

The results show that there were significant differences on each characteristic selected for at least one influence dimension. When compared with non-influentials, the influential members of teaching staffs have more years of professional preparation, have more years of total and of within school experience, teach at higher grade levels and are more likely to be male than female. This suggests that the bases of influence are similar to those which might be expected in any professionally-staffed organization, namely, expertise insofar as this might

be indicated by preparation, experience, and grade level to which a teacher is assigned. Sex and age may appear significant only because they are related to these other characteristics; however, some cultural factors may also be operative.

The analysis designed to determine whether basically the same individuals were influential on the communication dimension, the three reliance dimensions, and the attributed influence dimension revealed that 56 of 79 possible influentials were influential on three or more task area dimensions. When generalized influentials were defined as those who were influential on four or more task areas, 40 individuals still met this criterion. Table 1 also reveals that generalized influentials differed from all other respondents on variables of age, experience, preparation, and grade level taught; no analysis was carried out for sex.

The location of administrative personnel within the influence structure was determined from the rank order of individuals for each school on each influence dimension. Every principal was found to be influencial on the reliance dimensions and on the attributed influence dimension. Every one of the 18 principals ranked first on the reliance as to discipline dimension, and in most cases the vice-principal ranked second. All but one of the principals ranked first on reliance as to interpretation of school policies while on reliance concerning teaching methods, 14 principals ranked first and 4 ranked second. Only two principals did not rank first on the attributed influence dimension.

Results for the general communication dimension were not as clearly related to administrative position. Of the 14 principals who were classed as being influential on this dimension, 10 ranked first, 3 ranked second, and one was fifth. Ten of 23 vice-principals were influential on this dimension. The results suggest that although principals are consistently high in influence on certain specific aspects of reliance and communication, they do not always hold the highest rank in the more generalized communication and influence structures.

As might be expected, the 18 principals were all classed as generalized influentials as were 16 of 23 vice-principals when generalized influentials were defined as those influential on four or more dimensions. If generalized influentials had been defined as those who were influential and all five dimensions, 14 principals and 6 vice-principals would still have qualified.

The analysis suggests some similarities in the composition of influence structures across schools; designation as principal or vice-principal, and the possession of various professional characteristics are highly related to actual or potential influence. The results of this study are similar to those reported by House (1966) in his intensive analysis of one secondary school. This suggests that further analysis may verify

the possibility of generalizing across schools about influence structure characteristics which may appear initially to be highly specific to a particular school.

Subgroup Structures

The distribution of influence and the flow of communication are probably closely related to the subgroup structures within an organization. Although there has been considerable speculation about the configuration of the social structure of schools, there is little research evidence on the formation and characteristics of subgroups within schools. Through factor analytic procedures and the inspection of subgroups, this study attempted to provide some of the basic information which appears to be lacking in research literature.

Characteristics of subgroup structures. The analysis revealed considerable variation across schools in terms of the number of subgroups, the size of subgroups, and the relationships between subgroups. The following descriptions were developed in an attempt to identify different types of structures:

a. one generalized subgroup, few isolates, interconnections among all or nearly all members, no focal person.
b. one generalized subgroup in which communication or interaction focused on one particular person.
c. a structure consisting of two or more subgroups in which one is clearly separated from the others.
d. a structure involving two or more subgroups with little overlap in membership.
e. a structure of two or more subgroups with considerable overlap in membership.

The attempt to develop a typology was not highly successful since it yielded few pure types; however, even combinations of these descriptions proved to be useful in outlining the characteristics of the social structure of schools.

The analysis was carried out and descriptions were developed for both the communication and the socialization dimensions. Table 2 includes the classification of schools (identified by code number) by subgroup characteristics. Schools which are described by more than one characteristic are identified. For the schools in this sample, the most frequently occurring structures were two or more groups with either some or considerable overlap in membership. It is interesting to note that the first two characteristics are almost exclusively descriptive of communication and not of socialization structures while the fourth type appears to be repeated within the two dimensions for the same schools.

Number and size of subgroups. The number of subgroups in each school and the size of each subgroup were also determined. Table 3

TABLE 2

SCHOOLS CLASSIFIED BY SUBGROUP CHARACTERISTICS ON THE COMMUNICATION AND SOCIALIZATION DIMENSIONS

Subgroups Characteristics	Dimension	
	Communication	Socialization
1. One generalized subgroup--no focal person	4, 15*	18
2. One generalized subgroup with focal person	5*, 7, 8 13*, 15*, 18*	-
3. Two or more subgroups--one isolated	1*, 6*, 11*	1*, 5*, 6, 8, 17*
4. Two or more subgroups--little overlap	1*, 2, 3, 9, 11*, 13*, 14	1*, 2, 3, 10, 11, 14, 15, 16, 17*
5. Two or more subgroups--considerable overlap	5*, 6*, 10, 12 16, 17, 18*	4, 5*, 7, 9, 12, 13

*Each of these schools classified in more than one way.

shows the distribution of 51 communication subgroups and 57 socialization subgroups throughout the 18 schools in the sample. A single communication subgroup was identified in only two schools and a single socialization subgroup identified in only one school. Ten schools had two or fewer communication subgroups while only five had two or fewer socialization subgroups.

As might be expected, the number of subgroups within a school was related significantly to the number of staff members; the larger schools included more subgroups. The rank order correlation coefficient between number of subgroups and number of staff members was .53 for the communication dimension and .56 for the socialization dimension.

A similar test revealed no relationship between the size of subgroups and the number of staff members in the school. The data in Table 4 reveal that for both dimensions more than one-half of the subgroups included six or fewer members. The extremely large groups appear to occur more frequently in the communication than in the socialization dimension.

TABLE 3

FREQUENCY DISTRIBUTION OF SCHOOLS BY NUMBER OF SUBGROUPS IN COMMUNICATION AND SOCIALIZATION DIMENSIONS

Number of Subgroups	Type of Subgroup	
	Communication	Socialization
1	2	1
2	8	4
3	5	6
4	0	5
5	1	2
6	1	0
7	1	0

TABLE 4

FREQUENCY DISTRIBUTION OF COMMUNICATION AND SOCIALIZATION SUBGROUPS BY NUMBER OF MEMBERS PER GROUP

Number of Members	Type of Subgroup	
	Communication	Socialization
3 or 4	13	15
5 or 6	14	17
7 or 8	8	9
9 or 10	2	9
11 or 12	5	4
13 or more	9	3
Totals	51	57

Subgroup characteristics. One further aspect of the analysis of social structure included in this study was determining whether the members of different subgroups in the same school differed on selected characteristics. Because of the relatively small number of individuals within subgroups in any one school, statistical tests were almost meaningless. Only in the case of the largest schools were any statistically significant differences observed. Table 5 presents the results from one school for illustrative purposes. This particular statistical analysis revealed significant differences among subgroups on such average charactertistics as age, total years of professional preparation, and grade level.

TABLE 5

COMPARISON OF SIX COMMUNICATION SUBGROUPS IN TERMS OF MEAN VALUES OF SIX CHARACTERISTICS

Subgroup Number	No. of Members	Mean Age*	Total Exp.	School Exp.	Years of Prep.*	Grade Level*	Sex[a]
1	12	26.92	6.92	4.75	3.25	9.58	1.08
2	5	35.60	14.20	3.60	4.40	10.00	1.40
3	3	41.67	14.67	5.33	2.00	3.67	2.00
4	7	27.86	10.14	7.14	3.29	10.57	1.00
5	3	36.00	12.00	6.33	2.00	6.33	2.00
6	3	53.33	25.00	10.67	2.00	1.33	2.00
Grand Mean		33.00	11.52	5.82	3.09	8.27	1.36

*Differences between means significant $p \leq .05$

[a]Male = 1.00; Female = 2.00

Examination of the characteristics of subgroups in schools where two or more subgroups were identified revealed a tendency for subgroups to differ markedly on most of the six characteristics. It appeared that women teachers, who also had greater number of years of teaching experience and a lower number of years of preparation and who were teaching at lower grade levels, interacted with each other on task related matters and also for informal socializing. Younger male staff members who had a greater number of years of preparation but fewer years of experience and who were teaching at higher grade levels also formed exclusive subgroups. In some other schools a subgroup composed of older males, with about the same experience as the female group but with more years of preparation, formed a third type of group. Approximately one-third of the communication subgroups and two-fifths of the socializaton subgroups could readily be assigned to these categories without much distortion.

Task and Non-Task Related Structures

An attempt was made to compare the degree of similarity between social structures more closely related to the formal task and those related less directly to the task by comparing communication and socialization structures. One approach in the analysis was to determine the percentage of members which communication and socialization subgroups held in common with subgroups of the other type of the

same or larger size in the same school. The number of subgroups across the 18 schools with percentage of common membership in certain ranges were then tallied and the results are reported in Table 6.

TABLE 6

FREQUENCY DISTRIBUTION OF SUBGROUPS IN COMMUNICATION AND SOCIALIZATION DIMENSIONS BY PERCENTAGE OF OVERLAP IN MEMBERSHIP WITH SUBGROUPS OF SAME OR LARGER SIZE IN OTHER DIMENSION

Percentage common membership	Subgroup Type	
	Communication	Socialization
100%	7	14
80 - 99	9	13
60 - 79	15	17
40 - 59	17	11
Under 40	16	9

An examination of the table reveals that socialization subgroups appear to be more completely embedded within communication subgroups of the same or larger size than is true for the reverse. Communication subgroups tend to go beyond the boundaries of particular social subgroups. Apparently communication networks in task related matters include individuals who are excluded from social relations which take place within cliques or subgroups. Further analysis to determine characteristics of those who are excluded might provide useful insights into the nature of the social structure of a school.

In an earlier analysis, Bezeau (1966) applied a different approach to studying the relationship between instrumental (task related) and expressive (non-task related) structures within the same sample of schools. Instead of focusing on subgroups, he compared interaction patterns of individuals on task related communication with those on informal socializing. He observed that variations in similarity were related to certain personal characteristics of staff members. For example, it was observed that females had more congruent instrumental and expressive interaction patterns than did males; that is, females tended to discuss task related matters with the same individuals with whom they interacted socially. To a greater extent, males tended to interact socially with a different set of individuals from those with whom they communicated on more task related matters. It was also observed that the more years of teacher education an individual had, the greater was the difference between his position in the instrumental structure

and his position in the expressive structure. The variables of sex and professional preparation are probably related; the male with more years of preparation is also likely to have supervisory responsibilities and is likely to interact with different groups socially and on task matters.

Conclusion

The results presented in this report suggest that the analysis of subgroup and influence structures in schools need not be restricted to case studies or to small schools. The research procedures and methods of analysis which have been used in restricted studies can be extended to samples of schools and tentative generalizations appear possible. The practicality of the techniques also suggest possibilities for studying changing organizational patterns and relationships in emerging forms of school organization.

It seems evident that many administrative and organizational problems could be better understood as a result of a more complete analysis of organizational structure. Sociometric based studies described in this report should be used to supplement studies based on the bureaucratic model. In addition, social structure characteristics and indices can be used in more theoretically based studies to test hypotheses in which organizational structure is either the dependent or independent variable (Wiens, 1968; Scharf, 1967). An early test of the usefulness of a comprehensive strategy would be a research effort which seeks to obtain information about both social and bureaucratic structural characteristics of schools and which analyzes the relationship between the two.

References

Bezeau, L. M. The instrumental-expressive dichotomy in school staffs. Unpublished Master's Thesis, The University of Alberta, 1966. Pp. 103-104.

Blocker, C. E., & McCabe, R. H. *Relationships between the informal organization and the curriculum in six junior colleges.* Austin, Texas, 1964. P. 107.

Blocker, C. E., McCabe, R. H., & Prendergast, A. J. *A method for the sociometric analysis of the informal organization within large work groups.* Austin, Texas, 1964. Pp: 27-29.

Boyan, N. J. A study of the formal and informal organization of a school faculty. Unpublished doctoral dissertation, Harvard Graduate School of Education, 1951.

Breitkreuz, E. A. An analysis of school influence and subgroup structures. Unpublished Master's Thesis, The University of Alberta, 1967. Pp. 45-80.

Glanzer, M., & Blaser, R. Techniques for the study of group structure and behavior: I. Analysis of structure. *Psychological Bulletin,* 1959, 56 (September), 317-332.

House, J. H. An analysis of interpersonal influence relations within a school organization. Unpublished doctoral dissertation, The University of Alberta, 1966. Pp. 237-249.

Kolesar, H. An empirical study of client alienation in the bureaucratic organization. Unpublished doctoral dissertation, The University of Alberta, 1967.

MacKay, D. A. An empirical study of bureaucratic dimensions and their relation to other characteristics of school organizations. Unpublished doctoral dissertation, The University of Alberta, 1964.

MacRae, D. Direct factor analysis of sociometric data. *Sociometry,* 1960, 23, (December), 360-371.

McCleary, L. A study of interpersonal influence within a school staff. Unpublished doctoral dissertation, The University of Illinois, 1957.

Scharf, M. P. An investigation of the relationship between the professional role orientation and the social structure of teacher groups. Unpublished doctoral dissertation, The University of Alberta, 1967.

Wiens, J. Attitudes, influence and innovativeness: An analysis of factors related to innovativeness in educational organizations. Unpublished doctoral dissertation, The University of Alberta, 1968.

Wright, B. & Evitts, M. S. Direct factor analysis in sociometry. *Sociometry,* 1951, 14 (March), 82-98.

Alberta Journal of Educational Research
Vol. XIV, No. 4, December, 1968

This study examines the relationships of academic achievement, athletics, club membership, and dating to the social status of Canadian high school sutdents. The findings suggest cross-cultural differences since high academic performance and high social status are positively related for this sample of Canadian students.

HENRY ZENTNER
The University of Calgary
and
ARNOLD R. PARR
The Ohio State University

Social Status in the High School: An Analysis of Some Related Variables[1]

The purpose of this study is to examine the relationships of a number of selected variables to the social status of high school students. The concept of social status is defined as a person's rank, standing, or position in a social system. A position of high social status would be one in which a person enjoys prestige and receives deference from others. Social status for high school students, according to Coleman (1961), amounts to recognition and respect from peers.

More specifically, this study compares the leading crowd members and other students in three senior Calgary high schools. Leading crowd membership is treated as a position of high social status in the social structure of the high school. The analysis initially focuses upon the factors associated with high social status and concludes in a cross-cultural examination of the data.

Previous Research

Previous research findings indicate that high social status among American high school students is primarily a function of participation in what has been called the fun subculture (Clark, 1962). More precisely, the empirical data suggest that the value system of the student social structure is strongly non-intellectual, if not in some respects anti-intellectual. In other words, the values which are associated with

1 The present study was supported by a grant in aid from the Alberta Advisory Committee on Educational Research.

high academic achievement and other intellectual endeavours are not as pervasive in the student social structure as those associated with athletics, membership in clubs, and success with members of the opposite sex. Evidence to support this conclusion is derived from sociological studies by Coleman (1961) and Gordon (1957).

On the basis of these previous research findings the following general hypothesis was formulated: *Participation in athletics, membership in clubs and frequent dating are positively associated with high social status for high school students, whereas high academic performance is either not related or negatively related to high social status.* The analysis in this study centers around this general hypothesis.

Methodology

The populations sampled in this study include pupils from three senior high schools in the city of Calgary. These three schools were selected on the basis of their location in residential areas characterized by high, average, and low socioeconomic status. Calgary, the chief urban and industrial center of southern Alberta, had a population of 276,975 persons at the time of this study.[2]

Limitations upon resources precluded surveying all of the students enroled in the three schools. The samples were drawn from each school so that students from Grades 10, 11, and 12 and students classified by the school authorities as high, average, and low achievers be more or less equally represented. These stipulations were substantially realized. The actual school classes selected for inclusion in the sample were chosen on the basis of convenient availability on the day the survey was conducted. The total sample includes 766 students, and these 766 students constitute 38 per cent of the total enrolment in the three schools.

Membership in the leading crowd of each school was determined by simply asking the students to name the members of the group which more or less ran things among the students. The sample of leading crowd members consists of those students who were part of the original sample and were named at least twice by other students as being leading crowd members. The relationship between the sample sizes of the leading crowds and the total number of students named as leading crowd members is reported in Table 1. The percentages in Table 1 show that for both males and females, the samples of the leading crowds include over 60 per cent of all students named as leading crowd members.[3]

---

2 The source of this population figure is the April, 1963 Civic Census. See *Municipal Manual: City of Calgary* (Calgary, Alberta: Compiled by the City Clerk, 1963). P. 75.

3 The sample size of the leading crowd is high relative to the sample size of the total enrolment because in one of the three schools all the students who were named as leading crowd members also were included in the original sample. This factor combined with leading crowd membership in the other schools being well represented on the original samples accounts for the relatively large leading crowd sample size.

TABLE 1

NUMBER OF STUDENTS NAMES AS LEADING CROWD MEMBERS AND SAMPLE SIZES OF LEADING CROWDS BY SEX

Sex	Number of Students Named as Leading Crowd Members N	Sample Sizes of Leading Crowds N	Per Cent
Male	73	49	67%
Female	105	67	64%
Total	178	116	65%

The data for this study were collected by means of two questionnaires, one for males and one for females. The questionnaires closely resemble in content those utilized by Coleman (1961, pp. 338-356).

The analysis of the data involves primarily the reporting of percentage distributions which facilitate comparisons between the leading crowd members and other students. Chi-square tests of association are utilized in further analysis of the data. The .05 level of significance is the level at which statistical associations are accepted as significant.

The Findings

Academic Performance and High Social Status

The indices of academic performance which are examined include I.Q.'s, quality of achievement as assessed by the school authorities, type of program of study, and amount of time spent doing homework.[4]

The I.Q.'s of the leading crowd members. The findings reveal that for both boys and girls, but especially for the boys, leading crowd membership is positively related to having high I.Q.s. The percentages in Table 2 show that students who have high I.Q.s are overrepresented in the leading crowds. Furthermore, those students who have average I.Q.s are underrepresented in the leading crowds.

Quality of achievement of the leading crowd members. The findings concerning the quality of achievement of the leading crowd members

[4] The policy within these three schools was to group students into homogeneous classes on the basis of academic achievement. Thus, certain classes were designated as either high, average, or low in academic achievement by the school authorities, and it is on this basis that we obtained a measurement of quality of academic achievement.

TABLE 2

DISTRIBUTION OF LEADING CROWD MEMBERS AND OTHER STUDENTS BY MEASURED I. Q. RATING

Response Category	Males[a]		Females[b]	
	LCM	OS	LCM	OS
High Ratings	69%	49%	55%	49%
Average Ratings	29%	43%	39%	43%
Low Ratings	2%	4%	3%	3%
No Answer	0%	4%	3%	5%
Total	100%	100%	100%	100%
N	49	397	67	253

[a] $x^2 = 6.13$, df = 2, $p \leq 0.05$

[b] $x^2 = 0.62$, df = 2, not significant

relative to the other students are quite similar to the findings about the I.Q.s of the leading crowd members. The percentages in Table 3 show that students who are high achievers are overrepresented and those who are low achievers are underrepresented in the leading crowds. In other words, for both boys and girls the situation is that the leading crowd members are much more likely to be high achievers than low achievers.

Programs of study of the leading crowd members. The findings reveal that leading crowd membership for boys is very strongly related to taking the matriculation pattern as a program of study. The percentages in Table 4 show that boys who are taking the matriculation pattern are overrepresented in the leading crowd.

Amount of time leading crowd members spend doing homework. For the boys the findings reveal leading crowd membership to be related positively to the average amount of time spent doing homework. The percentages in Table 5 show that boys who do two hours or more homework a day are overrepresented in the leading crowds, and those who do only about one-half hour or less a day are underrepresented.

TABLE 3

DISTRIBUTION OF LEADING CROWD MEMBERS AND OTHER STUDENTS BY QUALITY OF ACHIEVEMENT AT SCHOOL

Response Category	Males[a]		Females[b]	
	LCM	OS	LCM	OS
High Achievement	59%	29%	63%	33%
Average Achievement	25%	34%	31%	33%
Low Achievement	16%	36%	6%	33%
No Answer	0%	1%	0%	1%
Total	100%	100%	100%	100%
N	49	397	67	253

[a] $x^2 = 17.93$, df = 2, $p \leq 0.05$

[b] $x^2 = 26.21$, df = 2, $p \leq 0.05$

TABLE 4

DISTRIBUTION OF RESPONSES OF LEADING CROWD MEMBERS AND OTHER STUDENTS TO THE QUESTION: "WHAT PROGRAM ARE YOU TAKING IN SCHOOL"

Response Category	Males[a]	
	LCM	OS
Matriculation Pattern	90%	66%
General	6%	22%
Vocational, Commercial, or Other	4%	12%
No Answer	0%	0%
Total	100%	100%
N	49	397

[a] $x^2 = 11.09$, df = 2, $p \leq 0.05$

Leading crowd membership and average amount of time spent doing homework are almost unrelated for the girls. As the percentages in the table show, the proportions of leading crowd members who do the various amounts of homework closely parallel the proportions of other female students who do similar amounts of homework.

TABLE 5

DISTRIBUTION OF THE LEADING CROWD MEMBERS AND OTHER STUDENTS BY AVERAGE AMOUNT OF TIME SPENT DOING HOMEWORK OUTSIDE SCHOOL

Response Category	Males[a]		Females[b]	
	LCM	OS	LCM	OS
About 1/2 Hour or Less	10%	35%	11%	12%
About 1 to 1 1/2 Hours a Day	53%	49%	55%	55%
About 2 Hours or More a Day	37%	16%	34%	32%
No Answer	0%	0%	0%	1%
Total	100%	100%	100%	100%
N	49	397	67	253

[a] $x^2 = 17.61$, df = 2, $p \leq 0.05$

[b] No Chi-square Statistic

Participation in the Fun Subculture and High Social Status

Participation in athletics, club and activity membership, and dating behaviour are analyzed as activities which characterize the fun subculture.

Athletics and leading crowd membership. The findings reveal that participation in athletics is related to leading crowd membership for the boys. The percentages in Table 6 show that football and basketball players are overrepresented in the leading crowd. This trend is stronger for football players than basketball players.

For the girls, being interested in athletics and having high social status are strongly associated. Those girls who display a high degree of interest in athletics are consistently overrepresented in the leading crowd. The percentages in Table 7 show that those girls who attended more than half of the home basketball games are overrepresented in the leading crowds. Furthermore, as the percentages in Table 8 show, those girls who show an uncommon interest in athletics by remembering

TABLE 6

DISTRIBUTION OF MALE LEADING CROWD MEMBERS AND OTHER STUDENTS BY WHETHER THEY PLAYED FOOTBALL AND/OR BASKETBALL AT SCHOOL DURING THE YEAR

Response Category	Football[a]		Basketball[b]	
	LCM	OS	LCM	OS
Yes	41%	23%	29%	15%
No	59%	76%	71%	78%
No Answer	0%	1%	0%	7%
Total	100%	100%	100%	100%
N	49	397	49	397

[a] $x^2 = 7.10$, df = 1, $p \leq 0.05$

[b] $x^2 = 4.30$, df = 1, $p \leq 0.05$

the basketball team's won-lost record are overrepresented in the leading crowds. Consequently, showing an interest in athletics is positively related to a girl's chances of being a leading crowd member.

TABLE 7

DISTRIBUTION OF FEMALE LEADING CROWD MEMBERS AND OTHER STUDENTS BY PROPORTION OF HOME BASKETBALL GAMES ATTENDED DURING THE YEAR

Response Category	LCM	OS
About Half or Less of Home Games	40%	78%
More than Half of Home Games	58%	19%
No Answer	2%	3%
Total	100%	100%
N	67	253

$x^2 = 39.69$, df = 1, $p \leq 0.05$

TABLE 8

DISTRIBUTION OF FEMALE LEADING CROWD MEMBERS AND OTHER STUDENTS BY WHETHER THEY REMEMBER THE BASKETBALL TEAM'S WON-LOST RECORD

Response Category	LCM	OS
Yes	46%	14%
No	52%	83%
No Answer	2%	3%
Total	100%	100%
N	67	253

$x^2 = 32.28$, df = 1, $p \leq 0.05$

Leading crowd members' participation in clubs and activities. Fraternity and sorority membership is strongly related to leading crowd membership. In fact, virtually all of the fraternity members are also leading crowd members as the percentages in Table 9 show. This table also shows that sorority members are highly overrepresented in the leading crowd.

TABLE 9

DISTRIBUTION OF LEADING CROWD MEMBERS AND OTHER STUDENTS BY FRATERNITY AND SORORITY MEMBERSHIP

Response Category	Fraternity Membership[a]		Sorority Membership[b]	
	LCM	OS	LCM	OS
Yes	49%	6%	49%	16%
No	48%	92%	46%	76%
No Answer	3%	2%	5%	8%
Total	100%	100%	100%	100%
N	33	302	67	253

[a] $x^2 = 62.25$, df = 1, $p \leq 0.05$

[b] $x^2 = 31.82$, df = 1, $p \leq 0.05$

As well as frequently being sorority members, the female members of the leading crowd tend to be those girls who have frequent or very frequent membership in clubs and activities in school. The percentages in Table 10 show that those female students who have frequent or very frequent membership in clubs and activities in school are over-represented in the leading crowds.

TABLE 10

DISTRIBUTION OF FEMALE LEADING CROWD MEMBERS AND OTHER STUDENTS BY FREQUENCY OF MEMBERSHIP IN CLUBS AND ACTIVITIES IN SCHOOL

Response Category	LCM	OS
Infrequent or No Club or Activity Membership	52%	77%
Frequent or Very Frequent Membership	48%	23%
No Answer	0%	0%
Total	100%	100%
N	67	253

$x^2 = 16.26$, df = 1, $p \leq 0.05$

Dating by leading crowd members. Frequent or very frequent dating is fairly widespread among leading crowd members compared to the other students, particularly for the girls. The percentages in Table 11 show that both male and female students who date frequently or very frequently are overrepresented in the leading crowds; going steady, on the other hand, is not strongly related to leading crowd membership. As the percentages in Table 12 show, boys and girls who go steady are slightly overrepresented in the leading crowds, but the results are not statistically significant.

Interpretation of the Findings

Rejection of the General Hypothesis

The main conclusion of this study is that the general hypothesis is not supported by the current findings. High academic performance is closely associated with high social status for high school students. Compared to the other students, the leading crowd members are more likely to have high I.Q.s and less likely to have average I.Q.s; furthermore, the leading crowd members are more likely to be high achievers than low achievers. As well, the boys in the leading crowds take the matriculation program more frequently than other students do, and a higher

TABLE 11

DISTRIBUTION OF LEADING CROWD MEMBERS OF OTHER STUDENTS BY FREQUENCY OF DATING

Response Category	Males[a]		Females[b]	
	LCM	OS	LCM	OS
Infrequently or Never Date	45%	65%	29%	41%
Frequently or Very Frequently Date	53%	32%	71%	56%
No Answer	2%	3%	0%	3%
Total	100%	100%	100%	100%
N	49	397	67	253

[a] $x^2 = 8.30$, df = 1, $p \leq 0.05$

[b] $x^2 = 3.90$, df = 1, $p \leq 0.05$

TABLE 12

DISTRIBUTION OF LEADING CROWD MEMBERS AND OTHER STUDENTS BY WHETHER THEY GO STEADY

Response Category	Males[a]		Females[b]	
	LCM	OS	LCM	OS
Yes	23%	16%	34%	29%
No	55%	59%	61%	67%
No Answer	22%	25%	5%	4%
Total	100%	100%	100%	100%
N	49	397	67	253

[a] $x^2 = 1.06$, df = 1, not significant

[b] $x^2 = 0.90$, df = 1, not significant

proportion of boys in the leading crowds than other students do two or more hours of homework a day.

There are factors other than academic performance which are associated with high social status. For boys, being an athlete, and for girls, a display of interest in athletics, is associated with high social status. Fraternity and sorority membership also is related to high social status. Active involvement in clubs and activities is also associated with high social status for girls. Frequent or very frequent dating is related to high social status for both boys and girls. To go steady, however, has little association with high social status.

Different Sources of High Social Status for American and Canadian Students

Since the general hypothesis which is not substantiated by the current findings is directly based on previous research evidence, it follows that the results of this study must be in at least partial disagreement with previous research findings. More precisely, the current findings do not tend to support certain conclusions proposed by Coleman about social status in the student social structure. Athletics for boys in this study is not as strongly associated with high social status as Coleman's evidence suggests. Furthermore, high academic performance has a very pervasive position in the student social structure which tends to dispute Coleman's claim concerning the nonintellectual nature of the students' attitudes.

What accounts for the variation between the current findings and the situation as depicted by Coleman for American high school students? It should be pointed out that Coleman relies heavily upon indices which attempt to measure the students' orientations and attitudes toward intellectual pursuits to document the nonintellectual nature of the student social structure. In contrast, in this study measurements of actual academic performance are utilized. In other words, qualitatively different types of measurement indices of the academic dimension are used in these two studies. It is suggested that this factor might account for some of the variation in obtained results.

This, however, cannot be considered a completely adequate explanation since some of Coleman's findings indicate that the leading crowd members, relative to the other students, tend to be high achievers. There appears to be a discrepancy between the values which the American students profess concerning intellectual endeavours and actual academic accomplishments. Unless such a discrepancy does in fact exist, the implication is that Coleman's measurements of the students' values may not be completely reliable when viewed in a cross-cultural perspective. That is, perhaps the student social structure in the high school is not as uniformly and strongly non-intellectual in orientation as Coleman suggests.

On the basis of the findings of this study it can be argued that the student social structures in these three schools are positively oriented toward intellectual pursuits. If the students were generally negatively oriented to academic success, it would seem very unlikely that they would accord students who are high in academic performance the high social status which accompanies leading crowd membership. In every instance for both boys and girls, those students who were high achievers were overrepresented in the leading crowds. Surely, such would not be the case unless the students in general possess positive attitudes toward academic success and strive for high academic performance.

Another possible interpretation of the findings of this study in relation to Coleman's findings is based on the concept of cross-cultural variations. It is possible that the differences in findings appear because of cultural differences between the United States and Canada. This implication is worthy of further investigation. Downey's (1960) study of regional variations within the United States and between the United States and Alberta regarding the role of the school would suggest that such differences could exist.[5] Downey found that Canadians placed greater emphasis on knowledge, scholarly attitudes, creative skills, aesthetic appreciation, and morality as outcomes of schooling than did Americans. In contrast the Americans emphasized physical development, citizenship, patriotism, social skills, and family living much more than did Canadians. Given the kind of cross-cultural differences suggested above, it does not seem at all surprising that the sample of Canadian students is much more pervasively oriented to academic accomplishment and less involved in athletics than their American counterparts. It is postulated that the prescribed role of the school in Alberta has not to the same extent produced an atmosphere conducive to the development of student values which are characteristic of the fun subculture in American high schools. The nature of the role of the school in Alberta, it would appear, orients the students to intellectual pursuits and academic performance more readily than does the American school.

References

Clark, B. R. *Educating the expert society.* San Francisco: Chandler, 1962. Pp. 244-258.

Coleman, J. S. *The adolescent society.* New York: The Free Press of Glencoe, 1961. P. 143.

Downey, L. W. Regional variations in educational viewpoint. *The Alberta Journal of Educational Research,* VI (4), 195-199.

Gordon, C. W. *The social system of the high school.* Glencoe, Illinois: The Free Press, 1957.

[5] It has also been suggested that Canada is somewhat less egalitarian, less achievement oriented, and less individualistic than the United States. See Seymour M. Lipset, *The First New Nation* (New York: Basic Books, Inc., 1963). Pp. 248-273.

BOOK REVIEW

NEW LOOK AT EDUCATION
By: John Pfeiffer (New York: *The Odyssey Press, 1968*).

This little (162 pages) book carries the subtitle "Systems Analysis in our Schools and Colleges." The author, who has written several popularized scientific works (e.g. *The Changing Universe, Thinking Machine, From Galaxies to Man*), prepared this work with the cooperation of Educational Testing Service. It purports to be an introduction to systems analysis which, in the author's own words, "nonspecialists can understand and which provides a number of clear examples." Since this is what the author set out to accomplish, the book must stand or fall on these two criteria of simplicity and clarity.

As far as simplicity is concerned, this book does measure up as a useful introductory work. The author summarizes the systems approach to planning and decision making and provides a series of so-called models which describe the systems analysis process. These models, which are really flow charts depicting steps or phases in the process, are easily understood by anyone who is familiar with the classical description of the decision-making process.

On the second count, the book has some strengths, but also some serious weaknesses. There is a profusion of examples drawn from a number of systems analysis projects which have been implemented in schools and universities. Brief descriptions of these projects do serve to underline the growing importance of systems analysis as a technique in education, but their very brevity makes them unsuitable as basic learning materials in this important area of study. Moreover, because the projects used a wide variety of techniques all of which are comprehended under the umbrella term "systems analysis," the reader may have some difficulty in drawing any but the most abstract generalizations from the descriptions. That is to say, one can discern the major steps in the "systems analysis process" in each of the projects; however, having done so, one has not really seen much more than the fact that systems analysis is a rational process which seeks the best solution for a given problem. Anyone familiar with the reconstructed logic of the scientific method or with the classical description of decision-making will not find anything new in this level of generalization.

The reason for the confusion, which these examples may leave in one's mind, is to be found in the fact that systems analysis is based on an interdisciplinary approach. The philosophical insights of general systems theory, the work of economists, planners, cyberneticions, and the special skills of operations researchers, are all part of the "bag" of the systems analyst. With this in mind, it should be understandable and

even forgivable when an introductory book suffers from lack of specificity with regard to the actual use of the techniques.

Two further points should be noted. There is the danger in using books of this kind that a reader may come away with the notion that he now "knows" what systems analysis is. Like every work which simplifies complex subject areas for the "millions," this one could stand in the way of a thorough, albeit demanding, study of the concepts and procedures associated with systems analysis. Finally, the book suffers from what must be termed a fatal weakness in an introductory work. There is no bibliography and no references are provided for the reader whose interest in this area is aroused for the first time.

In spite of its weaknesses, this book, if properly used, can serve to arouse the interest of both practitioners and researchers. If interest is aroused, the libraries of our universities abound with books and journal articles which can provide a solid basis in this relatively new, but already important, field of study.

D. A. MacKay
The University of Alberta

THE WORLD EDUCATIONAL CRISIS: A SYSTEMS ANALYSIS

By: Philip H. Coombs (Toronto: Oxford University Press, 1968)

The crisis referred to in the title is that since the 1950's, although student enrolments have doubled in many countries and educational expenditures have more than doubled, there is an *increase* in the number of adult illiterates in the world. Coombs attributes this crisis to three major causes: scarcity of resources (building, equipment, teachers), conservatism in education (an unwillingness to change), and the unwillingness of society to adapt (as when developing countries which need technical and vocational education attach undue prestige to academic and professional preparation). He proposes systems analysis as a framework for diagnosis and strategy.

Most readers will be familiar with the general outlines of systems analysis. In Coombs' presentation the educational process is thought of as a system, whose resource inputs are aims and priorities, students, management, structure and time schedule, content, teachers, learning aids, facilities, technology, quality controls, research, and costs. The outputs are more difficult to measure. A crude index of output is the number of students who complete various phases. A much more pertinent, but vastly more elusive output is the attitudes, values, and motivations of students. Still another, probably still more elusive, is the number and kind of leaders and innovators. The manpower output is im-

portant in the economy: are there enough people with the proper preparation to perform the work of society? Other outputs are research, to increase the store of knowledge and to solve problems, and international understanding.

Many economists talk in terms of systems and "input — output", and in this day and age, when in Canada education absorbs 8% of the Gross National Product, the economists have shown ever increasing interest in education and its costs! Coombs does not fall into the trap which others seem to. He recognizes the complexities, and refuses to select only the simple, measurable inputs and outputs. While this complicates a systems analysis of education, it avoids a biased and warped presentation. One of the reactions one gets from reading this book is the fairness and balance in Coombs' views. He refuses to oversimplify the complex.

An illustration of this fairness and balance is found in his discussion of money as an educational input.

> The questions about financial resources posed above cannot be divorced from their environment; education is but one part of a seamless web of things that make up life in a society. At any given time, the society's economy has just so much income to deploy. The amount that goes to education is subtracted from the amounts available for other purposes. For these reasons, the claims of education on national resources encounter the competitive counter claims of important material needs, such as investment in agriculture and industry, roads and housing, and important social needs, such as health, old age security, and unemployment relief. Regrettable, education's toughest adversary in more than a few countries is the military budget. But education even divides against itself in a competition for resources—where the rivals may be primary education versus secondary, or secondary versus university; expanded teacher training versus expanded construction of classrooms; and, of special importance, formal education versus nonformal.
>
> This competition demands the formulation of an order of national priorities, whether explicitly or implicitly. The settling of national priorities, however, especially where equally strong arguments produce a deadlock in reason itself, is a notoriously painful affair. The matter can be greatly aided by solid facts, rationally analyzed. But in the end the priorities are finally set, not by a planner's slide rule, but by a political process—a process sometimes marked by rough-and-tumble budget battles among ministries, or between them and legislatures, or within legislatures. What finally emerges usually reflects a blend of the values of the society, and the comparative strengths and strategies of contending pressure groups. Indeed, precisely on this account, it is important for educational leaders to master not only their own field but the language and techniques of economists as well, in order to be better armed for the defense of their own proposals in the annual 'battle of the budget.' Good rhetoric is no substitute for facts and analysis in winning these battles.

More than half of the book is devoted to the input-output systems analysis, but it also includes lengthy sections on nonformal education, international cooperation, and conclusions for strategy. A few of these latter are: modernized management, modernized teachers, modernized learning process, the strengthening of education finance, greater emphasis on nonformal education, and international cooperation. These, Coombs says, will help solve the world educational crisis.

The manuscript was originally prepared as a paper to be presented at the President's Invitational Conference on the World Crisis in Education (Williamsburg, Virginia, October, 1967). In its mimeographed form it had 212 pages with a technical annex of an additional 69 pages. The preface to the original version (but not included in the book) notes that the manuscript was prepared by invitation following a meeting of "eighteen distinguished educational leaders, from eighteen different countries." This preface emphasizes that the presentation is an *analysis,* not a statement of *policy.* The author, as director, generously acknowledges the contributions made by the staff of the International Insitute for Educational Planning, a UNESCO organization located in Paris. The book closely follows the original paper.

Readers should be warned that the title "*World* Educational Crisis" correctly describes the breadth and vision included in Coombs' treatment of the topic.

S. C. T. Clarke
The Alberta Teachers' Association

NOTES ON CONTRIBUTORS

ROBERT D. ARMSTRONG holds bachelor's and master's degrees from The University of Alberta and an Ed.D. from the University of California at Berkeley. He was formerly a teacher and principal in elementary schools of Alberta and is presently an Associate Professor in Elementary Education at The University of Alberta. He is the senior editor of a monograph on Teacher Education, a report of the invitational conference sponsored by the Department of Elementary Education.

ELMER A. BREITKREUZ is now under a three year contract teaching in Ghana. He obtained his M.Ed. in Administration at The University of Alberta.

DAVID J. CHABASSOL is an Associate Professor in the Faculty of Education at the University of Victoria. He is the author of several articles dealing with variables relating to teaching success. This year he is on sabbatical, gathering data on various aspects of adolescent development and attitude formation.

S. C. T. CLARKE was a professor from 1948 to 1958. For the last eight of these years he was in the Department of Educational Psychology at The University of Alberta. From 1958 to 1968 he was Executive Secretary of the Alberta Teachers' Association. He has rejoined the staff at The University of Alberta and will be Director of Summer Session and Evening Credit programs and professor of Educational Psychology.

LORENE EVERETT received her B.Ed. and M.Ed. degrees from The University of Alberta. She has taught in Alberta and with the Department of National Defence in Germany. She is presently lecturing at The University of Alberta in the area of Elementary Language.

WERNER W. LIEDTKE received his M.Ed. in Elementary Education from The University of Alberta. He is continuing his studies at The University of Alberta. He has had articles published in MATHEX (Encyclopedia Britannica of Canada).

D. A. MACKAY is Associate Professor of Educational Administration at The University of Alberta. His previous publications include articles in the *Alberta Journal of Educational Research, Canadian Education and Research Digest* and *Canadian Administrator.*

E. MIKLOS, currently on sabbatical leave from his position as Professor of Educational Administration at The University of Alberta, holds an appointment as Visiting Professor at the University of California,

Irvine. In addition to university teaching he has participated in activities of the Council on School Administration and served as Director of the Leadership Course for School Principals for three years. Recent publications include articles in the *CSA Bulletin* and the *Canadian Administrator,* and a chapter in a forthcoming publication entitled *Educational Administration: International Perspectives.*

L. D. NELSON obtained his Ph.D. from the University of Minnesota after graduating from The University of Alberta. He is co-editor of MATHEX published by Encyclopedia Britannica of Canada. He is a member of the Publications Committee and a member of the Editorial Board for a yearbook on the history of mathematics education, National Council of Teachers of Mathematics.

ARNOLD R. PARR is a Research Fellow at the Disaster Research Centre at the Ohio State University. He has published previously in the *Alberta Journal of Educational Research.*

DAVID C. THOMAS is a graduate of the Universities of British Columbia and Oregon. Formerly on staff at the University of Victoria, he is now Chairman of the Division of Education and Psychology at the University of Ottawa, Kansas. He has co-authored a previous study on factors related to teaching competence and has previously been published in the *Alberta Journal of Educational Research.*

H. ZENTNER is a Professor of Sociology, the University of Calgary. His publications include articles in the *Alberta Journal of Educational Research, The American Sociological Review, Sociology and Social Research, Sociological Inquiry* and *Anthropologica.*

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Table of Contents—Volume XIV, 1968